ASPHALT ROAD POTHOLE IMAGE DETECTION USING DISCRETE WAVELET TRANSFORM

BY

OYINBO, Adebayo Matthew
MENG/SEET/2017/7501

A THESIS SUBMITTED TO THE POSTGRADUATE SCHOOL FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA, NIGERIA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF ENGINEERING (M.ENG) IN COMMUNICATION ENGINEERING.

AUGUST, 2021

ABSTRACT

Potholes have been one of the major problems faced by road users globally, contributing to vehicular accidents on roads, quick wear and tear of vehicles, among others. This is a big menace to the economic growth of any nation. Overtime, advancement in vehicular technology and sensors, has led to the establishment of automated approach for detecting road anomalies. Though, these approaches can be categorized into 3D reconstruction-based approach, 2D vision-based approach and the vibrational-based approach. However, the 2D vision-based approach has received wide acceptance among the academia and the industry due to its intrinsic advantages over the 3D reconstruction-based and vibrational-based approaches. In this regard, this research work focuses on the detection of potholes on asphalt roads which is one of the prominent road anomalies, by improving on the detection accuracy of 2D vision-based approach, by presenting a 2D vision-based pothole detection algorithm based on discrete wavelet transform. A total of 400 road surface images were captured, pre-processed and segmented, using discrete wavelet transform and canny edge extractor. A deep learning approach using a pre-trained Convolutional Neural Network (CNN) was adopted for detection and classification of the segmented road images. The results obtained showed that this algorithm is able to detect potholes more accurately even during light illumination and foggy weather conditions, than proposed 2D vision-based techniques in literatures, having an overall accuracy of 93.33%, precision of 91.67%, and recall of 94.83%.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

ABBREVIATIONS

Acronyms	Meaning
CNN	Convolution Neural Network
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform
1D-DWT	One Dimension Discrete Wavelet Transform
1D-wavelet	One Dimension wavelet
2D-DWT	Two Dimension Discrete Wavelet Transform
JPEG	Joint Photographic Experts Group
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
ROC	Receiver Operating Characteristics
TPR	True Positive Rate
FPR	False Positive Rate
AUC	Area Under the ROC.
db	Daubechies Wavelets
LL	Low Low subband
LH	Low High subband
HL	High Low subband
HH	High High subband
JPEG	Joint Photographic Experts Group

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background to the Study

Road network plays important crucial role in any nation's economy (Aldagheiri, 2009), contributing to its economic growth and development with social benefits. Road networks provides access to employment, social, health and educational services (Fan & Chan-Kang, 2005; Walle, 1999). Roads open up more areas and stimulate economic and social development. For these reasons, road infrastructure is one of the most important of all public assets (Lemer, 1999). This is attributed to the fact that, it is one of the major means of transportation, that allow the vehicular movement of goods, services and human from one location to another (Tunde & Adeniyi, 2012). However, it is important to ensure that these roads are in good conditions at all times by continuous monitoring of its surfaces and repairs of areas with anomalies.

Road anomalies on asphalt road occur as a result of the road exceeding their maximum lifespan (Bello-Salau *et al*., 2014). The use of poor quality materials for construction, poor drainage system (Onoyan-Usina, 2013), excess road traffic and failure to comply with the standard road construction specification (Bello-Salau *et al*., 2014). These anomalies are usually observed in form of potholes, speed bumps and cracks (Akarsu *et al*., 2016). The adverse effects of these anomalies on roads cannot be over emphasized, ranging from discomforts experienced by drivers while plying such roads, vehicular damages, road traffic accidents, among others. Though, prevalent among these anomalies in developing nations is the potholes anomalies, which has contributed greatly to the rate of road traffic accident alongside over speeding (Ryu *et al*., 2015).

Several efforts have been made towards reducing the anomalous induced road traffic accidents among which include repair of anomalous road (Akagic *et al.*, 2017; Huidrom *et al*., 2013; Ouma & Hahn, 2016), ensuring compliance with acceptable road surface conditions, ensuring proper signage and drainage system as well as construction of new asphalt road surfaces.

Pothole has been one of the major problems faced by road users globally, and has been a contributing factor to vehicle accidents on roads, quick wear and tear of vehicles, among others. This is such a serious issue that several organizations and countries always try to show the effect of potholes on the economy at large. For example, the American Automobile Association estimated that about 16 million drivers had suffered damage from potholes in the last five years before the year 2016 within the United States alone (Byun *et al.,* 2018). Britain also made their estimation cost of fixing potholes on their roads to amount up to 12 billion pounds annually. While India has a record of over 3000 deaths in road accidents caused by potholes on their roads annually (Byun *et al*., 2018).

Nigeria is not left out, Federal Road Safety Corps (FRSC) of Nigeria on January of 2019, estimated that a total of 7,827 persons were involved in road accidents within a month ("540 killed in road crashes in January - FRSC - Premium Times Nigeria," 2019).

To reduce these menace and safe guard lives and properties on roads, most countries have set aside road maintenance agencies to occasionally check, repair and maintain road infrastructures across the country. This involve personnel's of these agencies, manually carrying out road surface inspections at various locations across the country for the detection of these road anomalies, which leads to waste of time, money, resources and man power.

In order to tackle these, there is a need of equipping vehicles with the capability of detecting and informing drivers and appropriate road maintenance agencies of the presence of these road anomalies. Furthermore, another approach is, also incorporating this technology into vehicular Ad hoc Network (VANETs) (Hegde *et al.,* 2014) as shown in Figure 1.1. Whereby, vehicles have the ability to detect and communicate with each other about possible road anomaly (potholes) encountered at different road locations and possible ways of avoiding such anomalies.

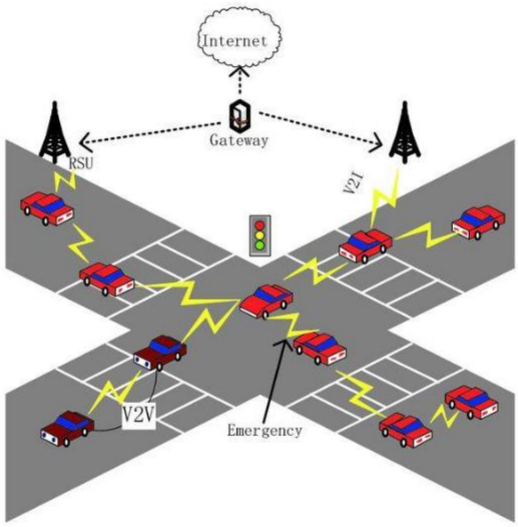


Figure 1.1: Vehicular Ad hoc Network (VANETs) (Wang, 2017)

1.2 Statement of the Research Problem

Manual Inspection of potholes for decades has been a method of detecting potholes and other road anomalies on asphalt roads in most part of the world. This method requires the deployment of personnel's to carry out this task which leads to waste of resources such as time, man power, and money. For the sake of safety of road users, there is a need to equip vehicles with the ability of detecting these road defects.

However, several research works emphasized on different automatic pothole detection techniques, this include Vibrational-based method, 2D Vision-based method and 3D reconstruction-based method (Kim & Ryu, 2014; Koch & Brilakis, 2011; Tsai & Chatterjee, 2017). The 2D Vision-based method however is preferred because of its cost effectiveness compared to 3D reconstruction-based and having higher accuracy compared to vibrational based method. Despite these claims, the 2D Vision-based method is still characterized by low detection accuracy during foggy weather conditions and at various light intensity according to literatures. In other words, the accuracy may fall at low light and extreme high light intensity conditions. To solve this anomaly, this research improves detection accuracy in the 2D Vision-based method.

1.3 Aim and Objectives

The aim of this research is to develop an asphalt road pothole image detection system using discrete wavelet transform. The objectives are:

1. To capture different road images, and pre-process the images using Discrete Wavelet Transform (DWT).
2. To segment and extract edges from the images using Canny edge extractor.
3. To develop a Convolutional Neural Network (CNN) based pothole detection system.
4. To evaluate the performance of the model using accuracy, precision and recall.

1.4 Justification of Study

Several research works have proposed different 2D vision-based road anomaly detection techniques. Koch & Brilakis (2011) proposed the use of morphological thinning and elliptic regression for detection, which has detection accuracy of 86%. Other 2D-vision based methods proposed by Akagic *et al.* (2017); Azhar *et al.* (2016); Ryu *et al*. (2015); Wang *et al*. (2017) all have improved detection accuracy than that of the technique proposed by Koch & Brilakis (2011), with the highest detection accuracy been 90.3% which was proposed by (Azhar *et al*., 2016).

However, the common limitation of the various 2D vision-based detection proposed is that, it is characterized by the low detection accuracy compared to that of 3D reconstruction-based approach (Tsai & Chatterjee, 2017). This is caused by the effect of different light illumination conditions on the road surface when images are captured. However the justification of this research work is that, there is room to increase the detection accuracy of 2D vision based techniques putting into consideration the effects of light. To achieve this, the research work employs discrete wavelet transform for the pre-processing of road images and a deep learning approach for detecting potholes on road images.

1.5 Scope of Study

This research work is restricted to the improvement of certain metrics in 2D Vision-based pothole detection. This metrics includes, accuracy, precision, and recall. To do this, captured road images were preprocessed using Discrete Wavelet Transform (DWT), segmented and inputted into a deep learning algorithm using a pre-trained Convolutional Neural Network (CNN).

1.6 Thesis Outline

This thesis is structured in five (5) chapters as follows: Chapter One is the Introduction, which introduces the need to automate potholes detection on asphalt roads. Chapter Two is the Literature review that presents various existing 2D vision-based road anomaly detection approaches. Chapter Three present Research Methodology and how the aim and objectives of this research was carried out. Chapter Four is the presentation of result and discussion, Chapter Five is Conclusion and Recommendations.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Pothole

A pothole is a depression on a road surface, usually asphalt pavement and can be a bowl shaped hole on pavement surface having minimum of at least 40mm depth. Potholes are holes in the roadway that vary in size and shape (Smith *et al.,* 2006). An example of a pothole image is shown in Plate I.

Plate I: Image of a Pothole

Potholes are generally caused when the pavement or the base cannot support the traffic loads. Two factors are always present in pothole failures: water and traffic. Heavy traffic or other factors may create cracks, which allow water to seep into the pavement base and soften it. The pounding of traffic causes the weak base to migrate, leaving nothing to support the pavement above it and thus initiating the formation of a pothole. Further traffic impact eventually causes the unsupported pavement to break up. Potholes may also be created in freeze/thaw situations. When water in the pavement or base freezes, it expands and pushes the pavement above it up. The swelling expansion forces can cause a pavement to weaken, resulting in pothole. Figure 2.1 shows how potholes on roads are formed.

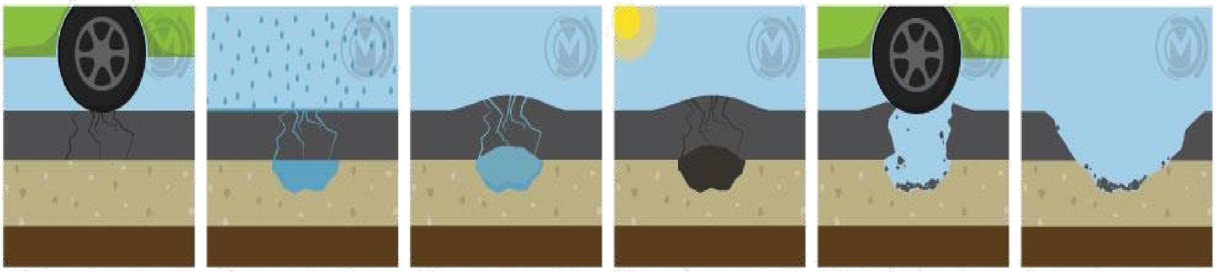


Figure 2.1 Formation of Pothole on Asphalt Pavement. (Smith *et al.,* 2006)

2.2 Pothole Detection

Various techniques are employed for pothole detection like manual notification using mobile applications, computer vision-based techniques, sensor-based techniques and so on (Aparna *et al.,* 2019). One of the simplest technique adopted for pothole detection involves the use of a mobile application such that whenever a user comes across any pothole on road, he/she snaps the picture of the pothole and sends it along with the location to the concerned authority so that appropriate action can be taken. Although simple, this technique requires human intervention and also depends on the voluntariness of the people. Thus, this technique may not prove to be of great use given the amount of money spent in developing the application and promoting it among people. Thus, automated pothole detection techniques are more useful. Pothole detection can be broadly classified into 3 types:

1. Vibrational-based methods (Kim & Ryu, 2014).
2. 3D reconstruction-based methods (Tsai & Chatterjee, 2017).
3. Vison-based methods (Koch & Brilakis, 2011).

Vibration-based methods is the use of accelerometer to measure and record the mechanical responses imposed by the road surface when driving through, and apply the responses to estimate the surface condition (Bello salau *et al.*, 2019; Zviedris, Mednis *et al.*, 2011).The measured signals from the accelerometer sensor is then processed using various processing techniques to estimate the severity of the road condition. Vibration-based method requires the vehicle to first make impact with the road anomaly (pothole) before being able to detect the anomaly, which is one of the major disadvantage of these method compared to the other detection methods. In general, vibration-based methods can be used in real-time processing, requiring small storage (Ryu *et al.*, 2015). However, it can only get a rough assessment of pavement potholes or even provide wrong results whereby, the joints of road can be detected as potholes, and potholes in the center of a lane cannot be detected using accelerometers due to no hit by any of the vehicle's wheels.

3D reconstruction-based methods: it makes use of 3D laser scanner methods, stereo vision methods by, and visualization using Microsoft Kinect sensor (Jahanshahi & Masri, 2012; Kim & Ryu, 2014; Tsai & Chatterjee, 2017). 3D reconstruction-based methods can obtain detailed information of the potholes more accurately than other detection methods. However, laser scanning systems cost too much to be applied at vehicle-level, and stereo vision methods need a high computational effort to reconstruct pavement surface through the procedure of matching feature points between different views. Although Microsoft Kinect sensor is cost-effective as compared to industrial lasers and cameras, the visualization method using Microsoft Kinect sensor is still a novel idea and further research is necessary (Kim & Ryu, 2014).

However the 2D vision-based detection method has proved to be an effective tool for automatic detection of potholes in roads. Due to the fact that manual pothole detection is labor intensive and time consuming, unsupervised and automatic pothole detection algorithms are being proposed in literatures under three (3) basic pothole detection classifications. However one of the most common limitation of the 2D vision-based detection approaches, is the inability to detect road anomalies accurately in foggy and light illumination conditions.

2.3 2D Vision-based Approaches

The use of 2D vision-based pothole detection approach method was proposed in Bubenikova *et al.* (2012), having advantages over 3D approach of high computational cost and equipment. To achieve detection, road image is enhanced using median filter to reduce the noise in the image. Then, image segmentation is implemented on the image, separating the defected and non-defected regions in the image using histogram shape-based thresholding. Due to geometric properties of a defect region, morphological thinning and elliptic regression were used to approximate the potential pothole shape. Subsequently, the texture of the potential defect shape is further extracted and compared with the texture of the surrounding non-defect region, so as to determine if the region of interest (ROI) represents a pothole. The strength of the proposed method in Bubenikova *et al*. (2012), recorded accuracy of 86%, precision of 82% and a recall of 86% in detecting road defects. However, additional visual characteristics, when extracting the pothole shape to improve pothole detection was not put into account, and the realization of the algorithm under different weather and light conditions (day/night) is still been tested. Furthermore, improving the detection accuracy rate, by taking additional visual characteristic into account when extracting the pothole shape, and also the use of machine learning to automatically train and classify potholes and non-potholes pavement textures, are all part of the future work proposed.

A robust vision-based detection method proposed in Huidrom *et al*. (2013) using video clip image processing techniques to automatically detect road defects. Collected road images were processed, to determine if the road contained distress or not. In order to achieve this, two algorithms were employed. First, image segmentation was carried out to separate the image into two different frames (frames with distress, and frames with no distress) using Distress Frames Selection (DFS) algorithm. After segmentation, Critical Distress Detection, Measurement and Classification (CDDMC) algorithm was then applied on the frames with distress to automatically detect and measure potholes, cracks and patches on the image. CDDMC algorithm makes use of image texture, shape factor and dimension for detection. However, the shadow of objects like trees captured in the road images can be wrongly detected as road defects.

A road anomalies detection method using morphological process was proposed in Buchinger & Silva (2014). Firstly, image segmentation is carried out on the collected images using Watershed algorithm, and morphological HMAX algorithm is then performed on the segmented image to detect road anomalies in the image. The result of the proposed method have a trade mark of simplicity in implementation, and it is able to detect cracks, but detect potholes more accurately. Out of 229 road defect images, the proposed method detected 211 accurately and 18 wrongly, but the detection accuracy of the method is affected by light and shadowing. The authors did not explicitly state to what degree this affected the detection accuracy but inferred this, because shadows of different objects reflected on the road surface in the images could be detected as a road defect.

Ryu *et al*. (2015) made use of histogram and closing operation of morphology filter for image segmentation. Candidate region extraction of potholes are extracted using features such as, size, ellipticity, linearity and compactness. Histogram Shaped-Based Thresholding was also applied to separate both the pothole region and a bright region, such as lane marking from the background region. Finally, Ordered Histogram Intersection (OHI) was used to decide if the image contained a road defect or not. The proposed method, had an accuracy of 73.5% with 80% precision and 73.3% recall in pothole detection. It also had a processing time of average of 46.8sec in processing 10 images. However, it has varying accuracy under different weather condition, vehicular vibration can also affect the detection accuracy. Also, potholes can be falsely detected according to the type of shadow and various shapes of potholes.

The detection of road patches proposed in Radopoulou & Brilakis (2015) made use of histogram equalization for image enhancement. The enhanced image is then converted into binary image using histogram shape-based thresholding algorithm, which allowed the separation of darker regions of the pavement image because, patches on pavement are usually darker than its surroundings. Morphological operation is then applied on the binary image. The proposed method made use of standard deviation of gray-level intensity values to describe texture for both a candidate patch and the healthy pavement around it. The method tackled the problem of reporting the same patch multiple times in a video sequence. It is also cost-effective and has fast processing time. Furthermore, the detection accuracy rate of the method is also affected by weather condition.

Theoretically, potholes are said to be elliptical in shape (Nienaber *et al*., 2015), but in reality, arbitrary shapes are possible due to irregular wear and tear of road surfaces. Potholes may either contain coarser (dry) or smooth (with water) texture, but the overall surface appearance of potholes can differ due to varying light illumination throughout the day. Due to this limitations which can affect detection accuracy, the proposed method in Azhar *et al*. (2016) made use of Histograms of oriented gradients (HOG) features to compute collected road images. HOG is based on the distribution of edge directions and cumulatively focused on the shape of an object. The collected image features are trained and then classified using Naïve Bayes classifier to label the image either as pothole or non-pothole image. The proposed method had a high accuracy of 90%, precision of 86.5% and recall of 94.1%.

A method based on fast adaptive approach for detecting road anomalies using computer vision was proposed in (Akarsu *et al*., 2016). The proposed method tried to adapt to various type of roads and detect defects on them. In other to achieve this, the different changes of road colors, had to be put into concentration. To do this, the color of the acquired road images are been controlled through customization, and the mean of the RGB image value is calculated and the image is ready for enhancement. To enhance the image, the obtained RGB image is converted to greyscale, and a blur version of the RGB image is also obtained. The combination of both grey and blur image helps to bring out the crack/ defect structure to the forefront. Binary transform and morphological operation is then carried out to get rid of noise in the image. After the image is enhanced, feature extraction and classification is carried out, and further details can be seen in (Akarsu *et al*., 2016). Experimental results showed that the proposed algorithm had high accuracy rates in different types of roads via customization, which have an average accuracy while running on all road types, and it has the ability to classify errors on different road surface. However, road color can affect the accuracy rates of the operation.

Using Wavelet Energy to separate potholes from non-potholes regions in road was proposed by (Wang *et al*., 2017). Although using wavelet energy field and morphological process can accurately detect potholes. To reduce the rate of false detection and increase pothole detection accuracy in an image, two processes (Wavelet energy field model and Markov Random field Model) were adopted to work hand in hand in the proposed method. After the enhancement and retriever of weak signal (potential pothole) among noise (surrounding areas) in the image, using Wavelet energy field by morphological processes and geometric judgment to detect pothole in the image. The detected pothole is then segmented using Markov random field model, where the original image is taken as the feature field, and the wavelet energy field is also taken as the label field. Morphological processing and edge extraction is performed after the pothole segmentation, making the pothole edge to be extracted accurately. The proposed technique is said to be better than that of Ryu *et al*. (2015) and Koch & Brilakis (2011), having an accuracy of 86.7%, with 83.3% precision and 87.5% recall. It also had good pothole segmentation results for different kinds of potholes, where about 88.6% of segmented potholes overlap degree is more than 0.85. However, the proposed techniques can be said to be complex, also having a processing time delay which is not suitable for real time implementation. Furthermore, the proposed technique accuracy is affected by light illumination conditions.

Due to the fact that most vision based pothole detection methods / algorithm are complex requiring much data for filtering and training, Akagic *et al*. (2017) proposed an algorithm to reduce this complexity by making use of RGB color space. Road image is been pre-processed, followed by image segmentation, to separate the asphalt pavement region from the rest regions in the image. The pothole region is extracted from the pavement image using the first, second and third level seed points, to narrow the search of pothole regions only. The pothole is detected by comparing the cropped images after Otsu thresholding is applied, and once all the linear and boundary shapes are eliminated in the image, the remaining regions are potential potholes. The proposed algorithm requires less data computation and can be implemented in real time, where the average execution time to detect a pothole from an image is 465.71ms with an average mean error of 7%. Also the accuracy of this method is said to be 82%, which is also suitable as pre-processing step for other supervised methods. However, the effectiveness of the method depends on the extraction of ROI accuracy and it is also affected by light illumination conditions.

Light illumination can affect the detection accuracy of defects (cracks, potholes) in roads, high contrast and low contrast both contributes to how accurate most defect detection algorithm could be. To solve part of this problem, Liu *et al*. (2017) proposed a method to deal with the effect of low contrast to increase the accuracy of detection. Image enhancement is applied on the road image gotten, combining multi-scale guided filter with gradient information, which helps to smooth, filter and preserve image edge details, which in turn enhances the defect while suppressing the noises. Dealing with the non-uniform illuminations in the image, adaptive threshold segmentation was adopted and the visual features which includes gradient and shape of the defect (crack), degree of area filling, the contrast between defects (crack) and background are used to refine the result. To automatically detect road cracks, a method was presented in D. Zhang *et al*. (2017), where Histogram equalization was used for image enhancement, and Bi – level thresholding which involves using two thresholds was applied on the crack like pixels of the improved image. A morphological close operation was applied on the image to connect the close pixels and morphological open operation was also used to delete isolated (noisy) pixels. After morphological operation, projection and detection is carried out, if distribution of dark regions is random, then, the inspected image do not contain defect but if inspected image contains a crack, at least, one or more very high peak will appear in projection results corresponding to alignment of crack pixels. The method has a crack image detection accuracy of 99%, and can be said to be simply in implementation, has fast processing time, were each image (4096 x 4096 pixels) have average of 100 milliseconds. However, the method's detection accuracy is also affected by light illumination conditions, and have high percentage of false alarm.

The use of superpixel segmentation for road anomaly detection was proposed in Sultani *et al*. (2017). To detect arbitrary shaped object of an input road image, the image is first divided into smaller consistent regions called superpixels. Superpixels also helps to preserve object boundaries. Various texture and intensity features such as Histogram of Oriented Gradients (HOG), Co-Occurrence Matrix (COOC), Mean Intensity (MI) and Intensity Histogram (IH), are all computed within each superpixel. Then, a support vector machine (SVM) classifier is trained for every feature separately in one-verse-all paradigm for object detection. Conditional Random field is then adopted to enforce superpixel neighbourhood label consistency, because the probabilistic scores of these superpixels are independently computed. Although the proposed method is complex and have higher processing time for detection, but, it is cost effective, efficiency is high, and can be used to detect any object on road surface.

Sharma *et al*. (2017) proposed a detection algorithm that makes use of Random forests classifier for Crack classification, by using selected binary image patches and ground truth binary class labels to train the system. To achieve the selected binary image, the road image was first pre-processed, were each RGB image patches are assigned a binary class label to each corresponding image pixels.

A mobile application which is able to recognize and highlight in real-time, three different categories of pavement distress (i.e. fatigue cracks, longitudinal and transversal cracks, and potholes) and their location was proposed in Tedeschi & Benedetto (2017). The application makes use of OpenCV local binary pattern (LBP) cascade classifier for training and classifying road distress. LBP features has immunity to a vast variety in lighting conditions making it suitable for the detection of pavement distresses. Input road images captured with the mobile phone are been processed with the application using LBP classifier to detect and classifier the road defects. The proposed method has low computational cost, detection accuracy of 72.8%, precision of 75.35% and recall of 76.95%. It also allows real time implementation. However, it required human intervention to carry out this detection.

The use of Fuzzy c-means clustering and morphological operations has also showed how effective they can be used for image processing. Due to this, a method of detecting potholes through image processing techniques making use of fuzzy clustering and morphological reconstruct was proposed in Ouma & Hahn (2017). First, input image is preprocessed and enhanced with median filter to reduce background noise in the image. To detect and recognize pothole in the image, the image is been passed through 2-D discrete wavelet transformation for more filtering. Adaptive soft- thresholding is further applied for detection of candidate defect pixels. Fuzzy c-mean clustering is used for recognition of regions of interest (pothole and non-defect surroundings). The proposed method had an average accuracy of 87.5% using the Dice coefficient of similarity, and average accuracy of 77.7% using Jac card index. However, the method is complex and detection accuracy is affected by lighting conditions and weather conditions.

Siriborvornratanakul (2018) proposed an algorithm for detecting road potholes with the use of on-board cameras on any vehicle. Images from the on-board cameras are first pre-processed using nonlocal means denoising algorithm, and histogram equalization is used to enhance contrast in the denoised image. For pothole detection, the equalized image is binarized using a threshold value, and morphological erosion is performed followed by dilation in order to remove tiny black noises and also connect the white area in the binarized image. Then, the contours in the morphed imaged is detected. The intensities of pixels inside each contour is measured, and the bigger the difference, the higher the first pothole-likelihood score. The proposed algorithm can successfully locate and pinpoints potholes despite unknown road surface's materials. However, the limitations of this proposed method includes, false alarms of pothole detection caused by non-pothole road damage, road patching and white line lane marking. Also, the algorithm is unable to detect large road distresses

Artificial Intelligence (AI), where machines are been trained and modeled to perform specific functions, has also proved to be a tool for detecting road defect and non-road defect. Due to this reason, a method of detecting road defects from road images using Artificial Intelligence was adopted in Hoang *et al*. (2018). Road images collected are been pre-processed using Gaussian filter for noise cancellation, steerable filter for image enhancement, and Integral projection which are all utilized for extracting features from the image. The proposed method, made use of two AI algorithms: Artificial Neural Network (ANN), which is a popular AI approach for pattern recognition, and least squares support vector machine (LS-SVM) are both been employed to generalize the decision boundary that classifies the instances of non-pothole and pothole classes.

An unsupervised road anomaly detection algorithm was presented in Buza *et al*. (2018), which make use of binary image and Simple Linear Iterative Clustering (SLIC) superpixels method for image segmentation .The application result of SLIC is fast and more memory efficient implementation of image segmentation. After the image pre-processing, the region of interest (ROI) of the processed image is extracted using different cluster selections. Otsu thresholding and image boundary shapes elimination are further applied on the image to carry out pothole detection. The detection accuracy of the algorithm presented is also affected by light illumination and weather conditions. Also, it did not detect smaller potholes with diameter less than 100 mm accurately.

Otsu thresholding is a good segmentation tool in image processing. Due to this reason, a crack detection method using Otsu thresholding was proposed in Akagic *et al*. (2018). In order to achieve this, image-slicing was carried out on the input image. The sub-images are then converted into gray scale images. The detection of cracks is based on the ratio between Otsu's thresholding and the maximum histogram value for every sub-image. After Otsu thresholding is performed on these sub-images, the sub-images are assembled back to form the resulting image.

Lee *et al*. (2018) proposed a method, which made use of both superpixel, wavelength and differential method for detection of potholes in images. Road Images gotten are been processed from RGB to grayscale. A superpixel is a clustering algorithm were a cluster of similar pixels in an image are formed. After the clustering process, it makes the outlines of cells which are clustered pixels. Wavelet energy field is used to show the texture of the image. So, it is used to check the cell that contains pothole or not. If the cell includes pothole, the wavelet energy field reveals the pothole by highlighting the cell.

Spectral clustering can be used to identify regions in an image. An algorithm making use of spectral clustering for pothole detection was presented in Buza *et al*. (2013). The algorithm made use of histogram-based thresholding for image segmentation. Otsu's image thresholding is further used to remove noise from the image to accurately distinguish objects in the image from the background. After segmentation, linear shapes and regions are removed. Spectral clustering is then used to identify the pothole regions in the image.

The proposed method in Xia (2018) made use of deep convolution networks for road distress detection on images. Features extraction of different road defects images are been trained on Conventional Neural Network (CNN) to detect any road defect from an image. The proposed method have the ability to detect all type of road defects, but it is complex, and the detection accuracy can be affected by weather condition.

The use of thermal imaging technology was adopted in Aparna *et al*. (2019). A thermographic camera, also called an infrared camera or thermal imaging camera is a device that forms an image using infrared radiation. This thermal imaging camera, differs from common cameras that forms an image using visible light, because it has the ability of capturing images at any weather conditions. The thermal imaging is used to capture road images and then, convolutional neural networks (CNN) is used to predict the presence of potholes from the road thermal images gotten. The proposed algorithm has detection accuracy of 97.08% and the ability to detect potholes at night and foggy weather conditions. However, thermal imaging cameras are more expensive than normal cameras.

2.4 3D Reconstruction-based Approaches

Computer stereo vision is the extraction of 3D information from digital images, which produces accurate information about a certain image. To this avail, a pothole detection method was proposed in Zhang *et al*. (2014) using stereo vision. This is achieved by the use of disparity calculation from the 3D image gotten to compare the potential defect regions in the image with its surrounds. Detection of potholes is carried out after disparity calculations, where a surface fitting algorithm is used to estimate the road surface, and points lower than the road surface are said to be potholes. After the pothole 3D points are determined, relevant pixels in the disparity images are segmented, which can also enable the pothole areas to be labelled. The volume of the pothole can also be calculated using the pothole cloud. The proposed method can be used to determine the exact location and depth of the pothole, and the ROI is easily determined. However, it has limitations such as: complexity in implementation, high computation cost, high processing time, high maintenance cost, and false detection also due to error in the disparity calculations.

A method of detecting and classifying potholes in roads using 3D technology and watershed method is employed in Tsai & Chatterjee (2017). The 3D pavement data is passed through watershed algorithm for segmentation. After segmentation, top surface with dimension width of 150mm or more are considered as potholes and are detected. The proposed method provided high detection accuracy (having accuracy of 94.97%, precision of 90.80% and recall of 98.75%), and severity levels of each potholes can be assessed. However, it has limitation in complexity, high cost of implementation and cannot be implemented in real time, due to the high processing time.

2.5 The Combination of Both 2D Vision-based and 3D Reconstruction-based Approach

A road crack detection algorithm was proposed in Medina *et al*. (2014) using the combination of both vision-based approach and 3D reconstruction-based approach for road anomaly detection.

For vision-based approach, input image is first preprocessed and filtered using Gabor filters, then passed through AdaBoost algorithm which creates two different classifier that can detect transverse cracks and longitudinal cracks. While for the 3D reconstruction-based approach, makes use of morphological open to remove noisy pixels in the binarized image. Hough transform is applied. The results of the 3D images analysis is then used as two binary classifiers in the Adaboost algorithm to detect the cracks. The proposed method detect cracks in the pavement of the road combining visual appearance and geometrical information and have a detection accuracy of 96.6% for transverse cracks and 93.3% for longitudinal cracks. However, the method is complex, has high equipment cost (implementation cost), and high processing time compared to other methods. Table 2.1 in Appendix B, shows the summary of related works on road anomaly detection.

2.6 Median Filter

Median filtering is a nonlinear method used to remove noise from images (Stork, 2013). It is widely used as it is very effective at removing noise while preserving edges. It is particularly effective at removing salt and pepper type of noise. The median filter works by moving through the image pixel by pixel, replacing each value with the median value of neighboring pixels. The pattern of neighbors is called the "window", which slides, pixel by pixel over the entire image (Erkan *et al*., 2018). The median is calculated by first sorting all the pixel values from the window into numerical order, and then replacing the pixel being considered with the middle (median) pixel value.

Median filter considers each pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. Median filter is an order statistic filter which is also non-linear (George *et al*., 2018). It replaces the center pixel with a value equal to the median of all the pixels in window. So, using a median filter will help reduce least and highest intensity value pixels, generally represented by the impulse noise and so the picture clarity improves. Median filter is simple and provides a reasonable noise removal performance, it removes thin lines and blurs image details even at low noise densities (Stork, 2013).

2.7 Discrete Wavelet Transform

Images have smooth regions interrupted by edges and abrupt changes. Although, Fourier Transform is a powerful tool for data analysis, but it does not represent abrupt changes efficiently, because it represent data as a sum of sine waves, which is not localized in time or space. In order to accurately analyse signals and images that have abrupt changes, there is the need of using a new class function that is well localized in time and frequency, thus the use of wavelet is employed (Vosoughi & Shamsollahi, 2009).

A wavelet is a rapidly decaying oscillation that has zero mean. Unlike sinusoid which extend to infinity, wavelet exist for a finite duration. Scaling and Shifting are the two key concepts of wavelets. Scaling is said to be the process of stretching or shrinking a signal in time. A stretch wavelet captures slowly varying changes in a signal, while a shrink wavelet captures abrupt changes in a signal. Shifting, refers to delaying or advancing the onset of the wavel*et al*ong the length of the signal (Williams & Amaratunga, 1994).

Wavelet Transforms are grouped into two: Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). The key application of DWT are denoising and compression of signals and images, as it helps represent many natural occurring signals and images with fewer co-efficient. The wavelet transform has gained widespread acceptance in signal processing and image compression and also in edge detection (Vosoughi & Shamsollahi, 2009). Even the Joint Photographic Experts Group (JPEG) committee released its new image coding standard, JPEG-2000, which is been based upon Discrete Wavelet Transform. Wavelet transform decomposes a signal into a set of basic functions. These basis functions are called wavelets. Wavelets are obtained from a single prototype wavelet called mother wavelet by dilations and shifting. Discrete Wavelet Transform has been introduced as a highly efficient and flexible method for sub band decomposition of signals. The 2D-DWT is nowadays established as a key operation in image processing. It is multi-resolution analysis and it decomposes images into wavelet coefficients and scaling function. In Discrete Wavelet Transform, signal energy concentrates to specific wavelet coefficients. This characteristic is useful for compressing images (Vosoughi & Shamsollahi, 2009).

In DWT, a time-scale representation of the digital signal is obtained using digital filtering techniques. The signal to be analyzed is passed through filters with different cut-off frequencies at different scales. It is easy to implement and reduces the computation time and resources required (Vosoughi & Shamsollahi, 2009). A 2D-DWT can be seen as a 1D-wavelet scheme which transforms along the rows and then a 1D-wavelet transform along the columns. The 2D-DWT operates in a straight forward manner by inserting array transposition between the two 1D-DWT. The rows of the array are processed first with only one level of decomposition. This essentially divides the array into two vertical halves, with the first half storing the average coefficients, while the second vertical half stores the detail coefficients. This process is repeated again with the columns, resulting in four sub-bands within the array defined by filter output. Image consists of pixels that are arranged in two dimensional matrix, each pixel represents the digital equivalent of image intensity. In spatial domain adjacent pixel values are highly correlated and hence redundant. In order to compress images, these redundancies existing among pixels needs to be eliminated. DWT processor transforms the spatial domain pixels into frequency domain information that are represented in multiple sub-bands, representing different time scale and frequency points (Williams & Amaratunga, 1994).

2.8 Segmentation

Image segmentation refers to the process of partitioning a digital image into multiple segments i.e. set of pixels, pixels in a region are similar according to some homogeneity criteria such as color, intensity or texture, so as to locate and identify objects and boundaries in an image (A. Smith, 2010). Several segmentation methods have been proposed in the literature. The choice of a segmentation technique over another and the level of segmentation are decided by the particular type of image and characteristics of the problem being considered.

The four main approaches used for image segmentation are: Threshold Techniques, Edge Detection Techniques, Region-based techniques, Partial Differential Equation (PDE), Clustering Techniques, Segmentation Based on Artificial Neural Network and Multi-objective Image Segmentation. One of the most commonly used is the “Edge Detection”. Edge Detection is a technique in which the points where image brightness changes sharply or formally are identified. These points where the brightness of the image changes sharply are organized under line segments called edges (Smith, 2010).

Canny edge extractor which is an edge detection technique is adopted in this work due to its excellent ability of detecting and extracting edges in images especially in noisy state by applying thresholding method. It is preferable to other edge detectors like Sobel extractor, because of its good localization and response, immune to noisy environment, and also having an inbuilt Guessian filter for further image denoising (Kabade & Sangam, 2016).

2.8.1 Canny Edge Extractor

Edges are important features in an image since they represent significant local intensity changes. They provide important clues to separate regions within an object or to identify changes in illumination. The Canny edge detector is popular because it is the most optimal method of finding edges with good detection, good localization and single response to an edge. Canny edge extraction determines edges by an optimization process and proposed an approximation to the optimal detector as the maxima of gradient magnitude of a Gaussian-smoothed image (Bhatnagar, 2017).

The Canny edge extractor also uses Gaussian smoothing to remove noise present in images. This technique finds edges by separating noise from the image before finding the edges of the image. It does not affect the features of the edges of the image. The Canny edge detector is based on computing the squared gradient magnitude. Local maxima of the gradient magnitude that are above some threshold are then identified as edges. There are three criteria's for Canny edge detection (Kabade & Sangam, 2016), they are:

i. Detection of edge with low error rate: which means that, the detection should accurately catch as many edges shown in the image as possible
ii. Good Localization: The edge point detected from the operator should accurately localize on the center of the edge.

iii. Single Response: An edge in the image should only be marked once, and where possible, the noise in the image should not create false edges.

Edge detection using Canny edge detector works in the following steps (Bhatnagar, 2017), which are:

a. Smoothing: - The image is smoothened using the Gaussian filter, to reduce noise.
b. Finding Gradients: - Edge pixels are those where there is a sharp change in gray level values. These pixels are identified by computing the gradient of the image. The gradient is a unit vector which points in the direction of maximum intensity change. The vertical and horizontal components of the gradient are computed firstly and then the magnitude and the direction of the gradient is computed.
c. Non-Maxima Suppression: - Mainly edge thinning is performed in no-maxima suppression. In this step, on the basis of gradient magnitudes, the thick edges in the image are converted to approximately thin and sharp edges which can be further used for recognition purpose. In this process, the image is scanned along the edge direction and discards any pixel value that is not considered to be edge which will result in thin line in the output image.
d. Double Thresholding: - Two threshold values are considered in Canny edge detection technique, T1 = High Threshold, T2 = Low Threshold. The pixels having values of gray scale the pixels have values of gray scale level between T1 and T2, the result is depending on the neighboring pixels.
e. Edge Tracking by Hysteresis: - Edges that do not connect to a very certain (strong) edge are discarded in the final output image. Strong edges are interpreted as "Certain Edges" and are included in the final edge image. Edges that are not strong edges but are linked with strong edges are included in the output image. Level

higher than T1 are strong edge pixels, and the result is edge region. The pixels having values of gray.

2.9 Convolutional Neural Network (CNN)

Machine Learning (ML) algorithms belong to an area in Artificial Intelligence (AI), which endows intelligence to computers by learning the underlying relationships among the data and making decisions without being explicitly programmed. Different Machine Learning algorithms have been developed since the late 1990s, for the emulation of human sensory responses such as speech and vision, but they have generally failed to achieve human-level satisfaction. The challenging nature of machine vision tasks gives rise to a specialized class of Neural Networks, known as Convolutional Neural Network (CNN) (Yadav & Bethard, 2019).

CNNs are considered as one of the best techniques for learning image content and have shown state-of-the-art results on image recognition, segmentation, detection, and retrieval related tasks (Khan *et al*., 2019.). The success of CNN has captured attention beyond academia. In industry, companies such as Google, Microsoft, AT&T, NEC, and Facebook have developed active research groups for exploring new architectures of CNN (Deng *et al*., 2014). At present, most of the frontrunners of image processing competitions such as Microsoft Cognitive Toolkit, are employing deep CNN based models.

The topology of CNN is divided into multiple learning stages composed of a combination of the convolutional layer, non-linear processing units, and subsampling layers (Khan *et al*., 2019). Each layer performs multiple transformations using a bank of convolutional kernels (filters). Convolution operation extracts locally correlated features by dividing the image into small slices (similar to the retina of the human eye), making it capable of learning suitable features. Output of the convolutional kernels is assigned to non-linear processing units, which not only helps in learning abstraction but also embeds non-linearity in the feature space. This non-linearity generates different patterns of activations for different responses and thus facilitates in learning of semantic differences in images. Output of the non-linear function is usually followed by subsampling, which helps in summarizing the results and also makes the input invariant to geometrical distortions. The structure of CNN can be seen in Figure 2.2.

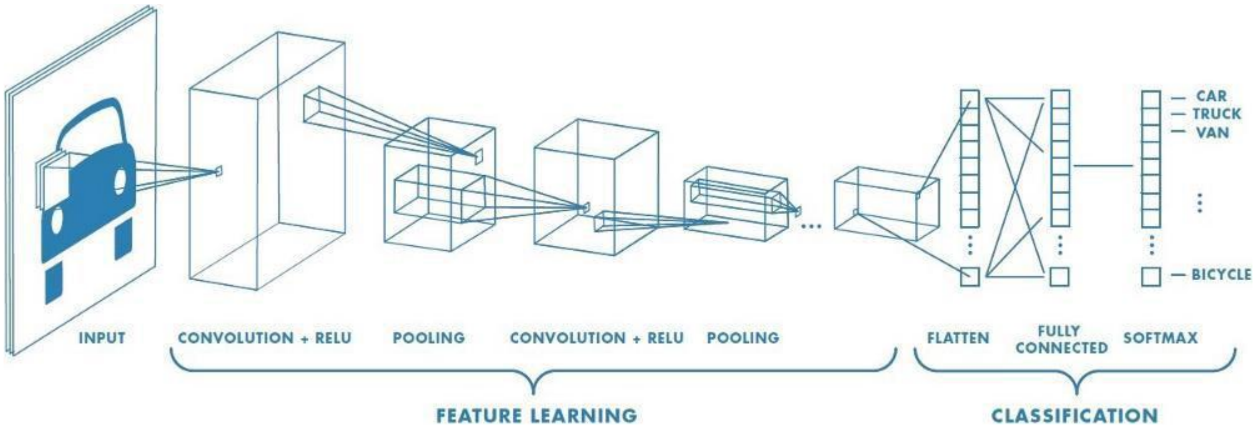


Figure 2.2: Structure of Convolutional Neural Network (CNN) (Gulcu & Kus, 2019)

There are several pre-trained convolutional neural networks, such as LeNet, AlexNet, VGGNet16, GoogleNet and ResNet (Khan *et al*.,2018). ResNet (Residual Neural Network) (He *et al*., 2016) stands out due to its higher accuracy and less error rate in network performance especially in image classification. Though there are different models of ResNet like, ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152 (Khan *et al*., 2019). However, ResNet50 has better performance (fewer error rates on detection tasks), compared to ResNet18 and ResNet34, and less complex, compared to ResNet101 and ResNet152, in the ResNet family (He *et al*., 2016).

2.9.1 ResNet50

ResNet50 is a convolutional neural network that is trained on more than a million images from the ImageNet database. The network is 50 layers deep and can classify images into 1000 object categories (He *et al*., 2016) as seen in Figure 2.3. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224-by-224.

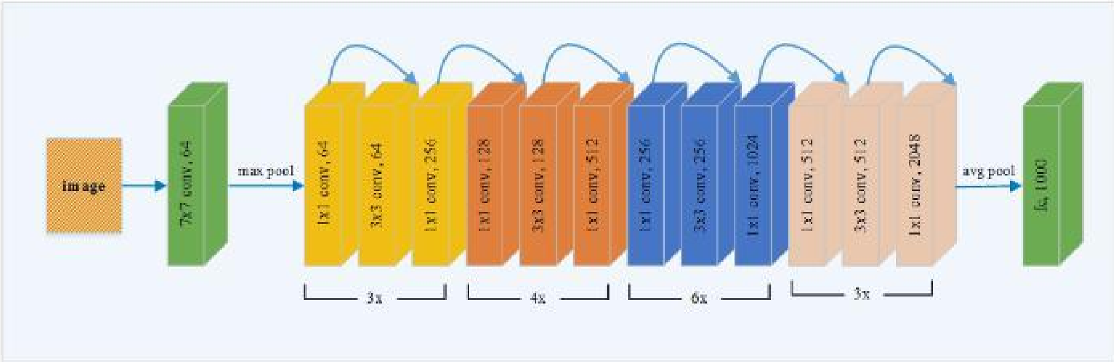


Figure 2.3: A Block Diagram Representation of Pre-trained Resnet-50 Architecture. (He *et al*., 2016)

CHAPTER THREE

3.0 MATERIALS AND METHODS

The methodology used for the realization of the research objectives and achieving the aim of the research is presented. This comprises of both the materials used for the research and the methods used for the implementation of the set objectives.

3.1 Materials

The following materials were used for the realization of the aim of the research:

1. A camera:

   An android mobile phone (Hauwai Y9) with a back camera of 13MP + 2MP, F/1.8 + F/2.4 aperture was used to capture road images for the dataset meant for this research work.

2. A laptop computer:

   A MATLAB (R2018b) prototype with Image Processing Toolbox was installed on an HP 2000 laptop with 6GB RAM and 500GB hard disk, running on Intel® Core(TM) i3-3110M CPU @ 2.40GHz processor in other to simulate the research algorithm.

3. A Global Position System Data lodger:

   A road surface profiler system developed by Bello-salau *et al*. (2018) was used to pin point and get the exact GPS co-ordinate of the road anomalies images captured on Google map. The profiler provides a map of different pothole points existing along the road.

3.2 Methods

This section presents the detailed techniques used for the realization of the research aim. This include the image data acquisition process, the pre-processing stage, the segmentation stage as well as the classification stage as summarized by the block diagram in Figure 3.1. An explanation of the process in each block is presented herewith.

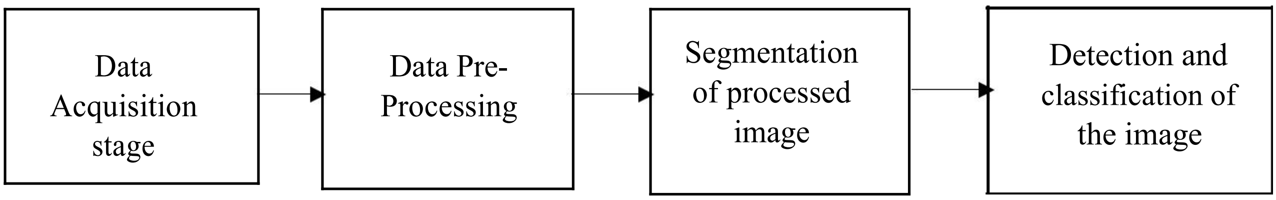


Figure 3.1: Block Diagram of the Research Methodology.

3.2.1 Data Acquisition

Due to the unavailability of image datasets characterized by different forms of road anomaly (pothole) and smooth roads in Nigeria, road images were manually captured to form dataset used in this research work. Note that a database containing road surface conditions was reported in (bellosalau *et al.*, 2018). However, the datasets contained in the database were only one-dimensional signals and not images. Thus, necessitating the need for data acquisition stage. A Hauwei Y9 android mobile phone camera with 13MP+2MP dual back camera was used for the image data capturing during drive test, because of its high resolution that ensures the clarity of the captured image dataset. These images, captured off the vehicle at different time intervals of the day are located at the Eastern by-pass road (Tundu wada north), Minna Niger state, Nigeria. To ensure high efficiency at different luminosity, images were captured as shown in Figure 3.2. Some were captured at the early hours of the morning (8am – 10am), before the sunset, others, when the intensity of the sun was high (1pm – 4pm). Also, some of the road images were captured after rainfall, in order to be able to get road images during the foggy weather in which potholes are often time filled with water. All the images were captured from the duration of 23rd October, 2019 to 27th October, 2019.

A total of 400 road surface images were captured to form the dataset. Half of it were characterized with potholes, while the remaining were good road surface images (including line markings).

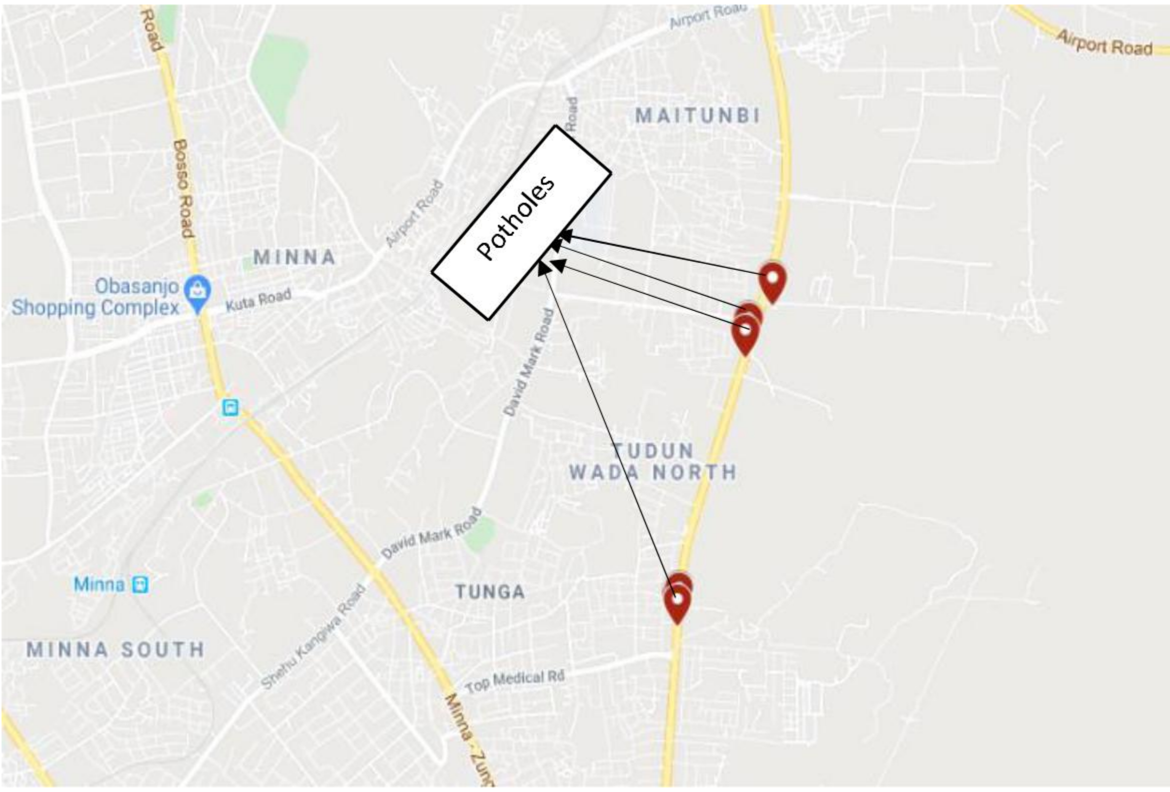


Figure 3.2: Google Map of Eastern by-pass Road, Minna Niger State, Nigeria. (Google, 2020)

3.2.2 Image Pre-Processing

MATLAB was used to simulate the entire image processing. During this work, the 2D vision-based pothole detection makes use of edge detection as a feature for detecting potholes within road images.

The road images in the dataset vary in sizes. To reduce the time taken to pre-process the images, all images were resized to have an image size of 400 * 300 pixel in the MATLAB prototype. After image resizing, grey scaling of the captured RGB images were done.

Afterwards, the images were passed through a 3 x 3 kernel median filter in order to denoise the captured images while preserving its edges.

The preserved edges of the image were enhanced by the application of a 2-D discrete wavelet transform (DWT) that decomposes the image into both the low pass and high pass filter component. The ‘Haar’ mother wavelet was used because of its high preserved edges of the filtered pothole images when compared to ‘dB’, ‘symlet’ and other dual basis functions. An approximate and detail coefficient of the inputted images were obtained from the low pass and high pass filters. The output of the approximate coefficient in the DWT LL sub-band, which contain more detail information was then used for the segmentation process.

3.2.3 Evaluation

Let the input image be f(x,y).

Applying median filter to the input image, therefore (George *et al.*, 2018):

$$g(x,y) \quad med\{f(x \quad i, y \quad j), i, j W \quad \}. \tag{3.1}$$

Where,

$g(x,y)$ is the image output of the median filter.

W is the two-dimension mask: the mask size is n x n (3 x 3).

Therefore, $g(x,y) \quad med\{f(x \quad i, y \quad j), i, j \quad (3 X 3)\}$ (3.2)

The output of the median filter $g(x,y)$ is fed into the DWT as shown in Figure 3.3.

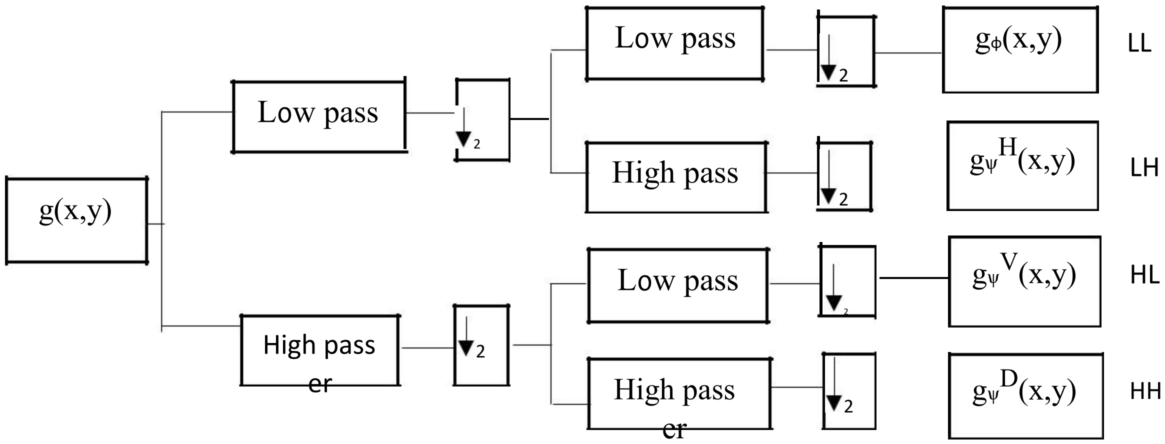


Figure 3.3: A 2-Dimensional Discrete Wavelet Transform.

Let the discrete wavelet transform function be f(x, y) of size (M * N).

Therefore, the approximate coefficient (LL sub-band) of the discrete wavelet transform can be expressed as (Williams & Amaratunga, 1994):

$$g(j_0, x, y) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)\, \varphi_{j_0, m, n}(x, y) \tag{3.3}$$

Where,

j_0 is an arbitrary starting scale.

$g(j_0, x, y)$ Coefficients define an approximation of f(x,y) at scale j_0.

3.2.4 Segmentation Process

In the segmentation process, after denoising and enhancement of edges of the images captured with the help of discrete wavelet transform, Canny edge extractor was further employed to segment and extract the edges of the image. Apart from having an inbuilt Gaussian filter for further image denoising, Canny edge extractor was used for segmentation because of its ability to detect and extract edges in images characterized with noise. During the simulation, it was noted that an experimental threshold value of

0.75 gave a better result for the image edge extraction. Figure 3.4 presents flowchart of the image pre-processing and segmentation process simulated in MATLAB towards an improved road anomaly detection and classification.

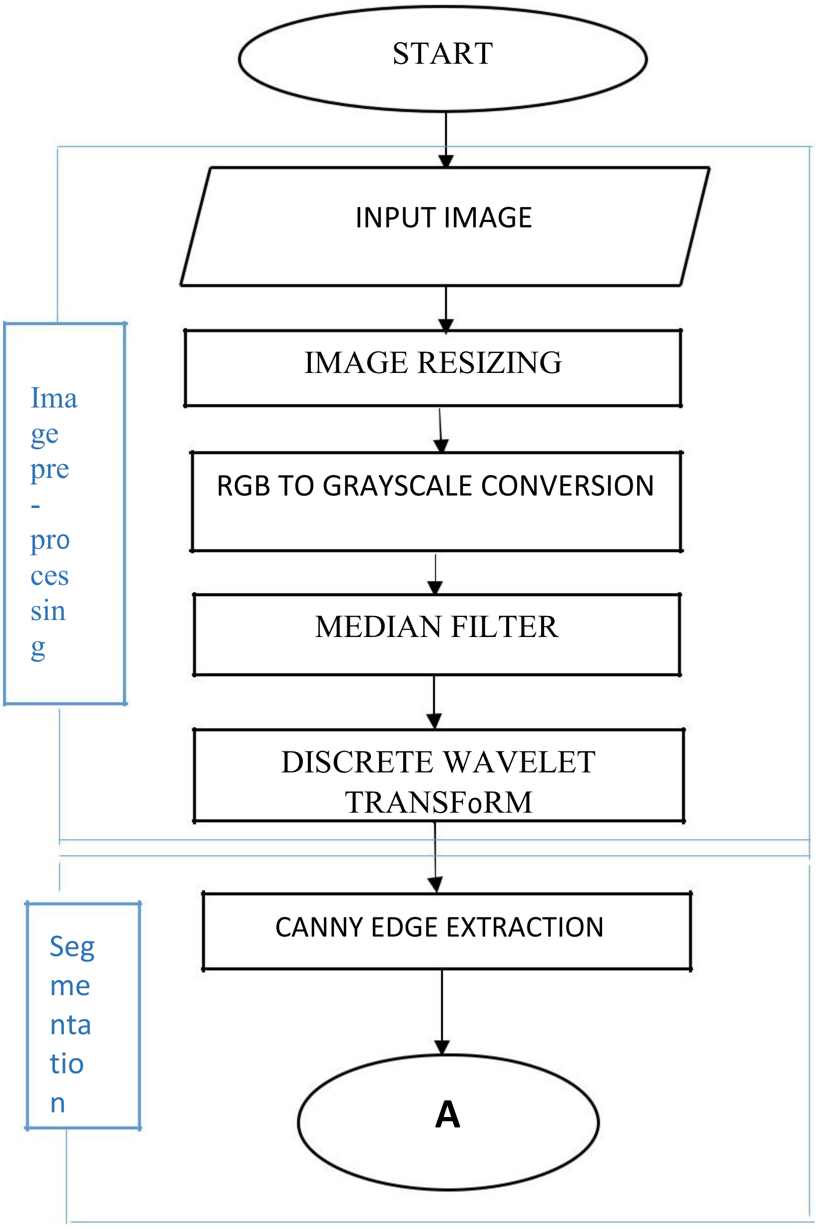


Figure 3.4: Flowchart of the Image Pre-processing and Segmentation Process.

Where, A is the output image after segmentation, which is fed to the classifier model.

3.2.5 Detection and Classification using CNN

In order to detect and classify the pre-processed and segmented road images, a deep learning approach using a pre-trained CNN known as ResNet50 which is majorly used for image classification, was adopted because of its better performance (fewer error rates on detection tasks), compared to ResNet18 and ResNet34, and less complex, compared to ResNet101 and ResNet152, in the ResNet family. The pre-trained CNN was also simulated on the MATLAB prototype by installing a ResNet50 module on the MATLAB IDE.

To realize the work, 70% of the dataset for training the CNN had to be created, by pre-processing and segmentation. As illustrated in Figure 3.4, the segmented road images are then grouped into two classes (pothole class and smooth road class) and kept in separate folders. Also, it should be noted that, segmented images with line marking edges were placed in the smooth road class folder. The dataset for training the CNN was then imported using the fullfile(); function on the MATLAB.

The ResNet50 module is loaded in the simulation and the training dataset was used to train the CNN (ResNet50 model). After training, the output "A" as shown in Figure 3.3 in the image pre-processing and segmentation algorithm was linked to the ResNet50 model. To test this algorithm, the remaining 30% road images from the dataset pre-processed and segmented was fed to the CNN. The CNN then detects and classified the inputted road images as either a pothole class or a smooth road class as shown in Figure 3.5.

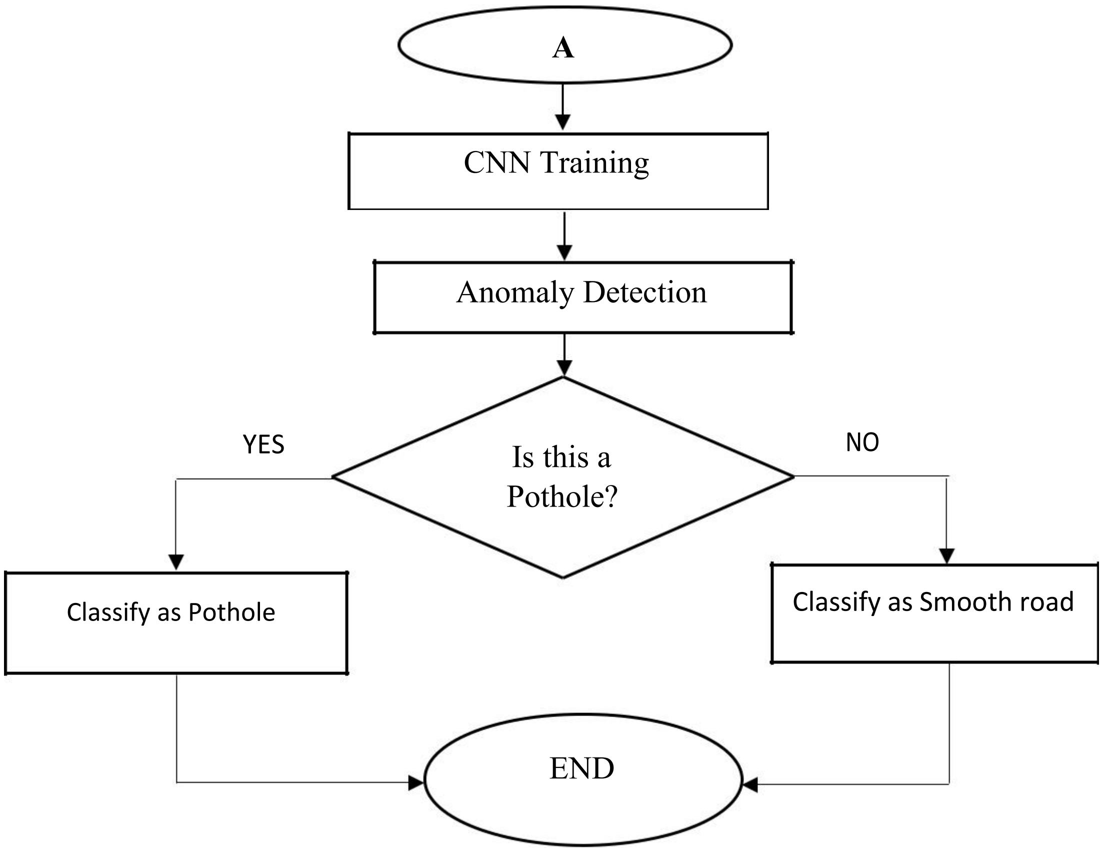


Figure 3.5: Flowchart of the CNN Process of Detection and Classification

To check the classification performance of this model, a Receiver Operating Characteristics (ROC) curve was used to visualize this performance. ROC curve is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters, i.e. True Positive Rate on the Y axis and False Positive Rate on the X axis.

3.2.6 Evaluation Process

To evaluate the performance of this research work and also compare the result obtained with other 2D vision-based methods in literature, a performance matrix used as a standard for validating the performance of the 2D vision-based pothole detection algorithms was deployed. In Figure 3.6, a confusion matrix, which is a specific table layout that shows the prediction results of a classification model in machine learning was used to interpret the detection and classification result of this research work. The result obtained from the confusion matrix was used to evaluate performance of the algorithm in terms of accuracy, precision and recall, and are all expressed in the equations (4), (5) and (6) respectively.

Accuracy can be described as the average correctness of a classification process, in other words, it means the degree to which the result of a measurement conforms to the correct value, and essentially refers to how close a measurement is to its agreed value.

Precision can be described as the detection exactness, showing how good a model is at whatever it predicted. For example, a model predicted 10 items and it turned out to be wrong at one, the precision is 9/10 (90%).

Recall is a measure for detection completeness, and can be referred to as, the ratio of the correct predictions and the total number of correct items in the set. For example, for a detector model, if there are 40 items and the model detect 20 items and they were correct, the recall is 50%. If it detected all 40, the detector can be said to have 100% recall.

An actual example involving precision and recall is having a detection model, if there are 20 items and the model predicted 10 items. Out of these, 4 items were predicted wrongly and the remaining 6 were correct. In this case, the precision of the model is 60% (6/10), while the recall is 30% (6/20).

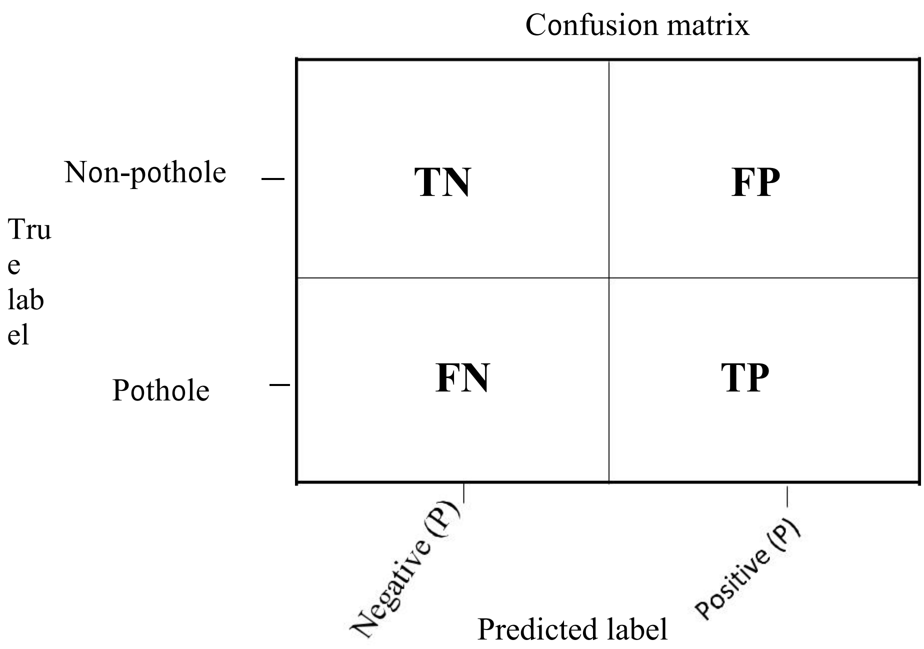


Figure 3.6: Confusion Matrix.

Where;

TP is the True Positive.

TN is True Negative.

FP is False Positive.

FN is False Negative.

Basically, for an already labeled image dataset, if a road anomaly exists being termed positive, and the algorithm was able to detect and classified the anomaly as potholes, then a True Positive (TP) is declared. But, if wrongly detected and classify as smooth road (non-pothole), then a false negative (FN) is declared. Furthermore, if the labeled image data sample is a smooth road with no anomaly, and the algorithm was able to detect and classify appropriately as smooth road (non-pothole), then a True Negative (TN) is declared, otherwise, a False Positive (FP) is declared.

Therefore,

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (3.4)$$

$$\text{Precision} = \frac{TP}{TP+FP} \qquad (3.5) \qquad \text{(Wang } et\ al.\text{, 2017)}$$

$$\text{Recall} = \frac{TP}{TP+FN} \qquad (3.6)$$

To plot the ROC curve, results obtained from the confusion matrix were also used to calculate the True Positive Rate (TPR), which is also called Sensitivity or Recall, expressed in equation (6). False Positive Rate, can be expressed as:

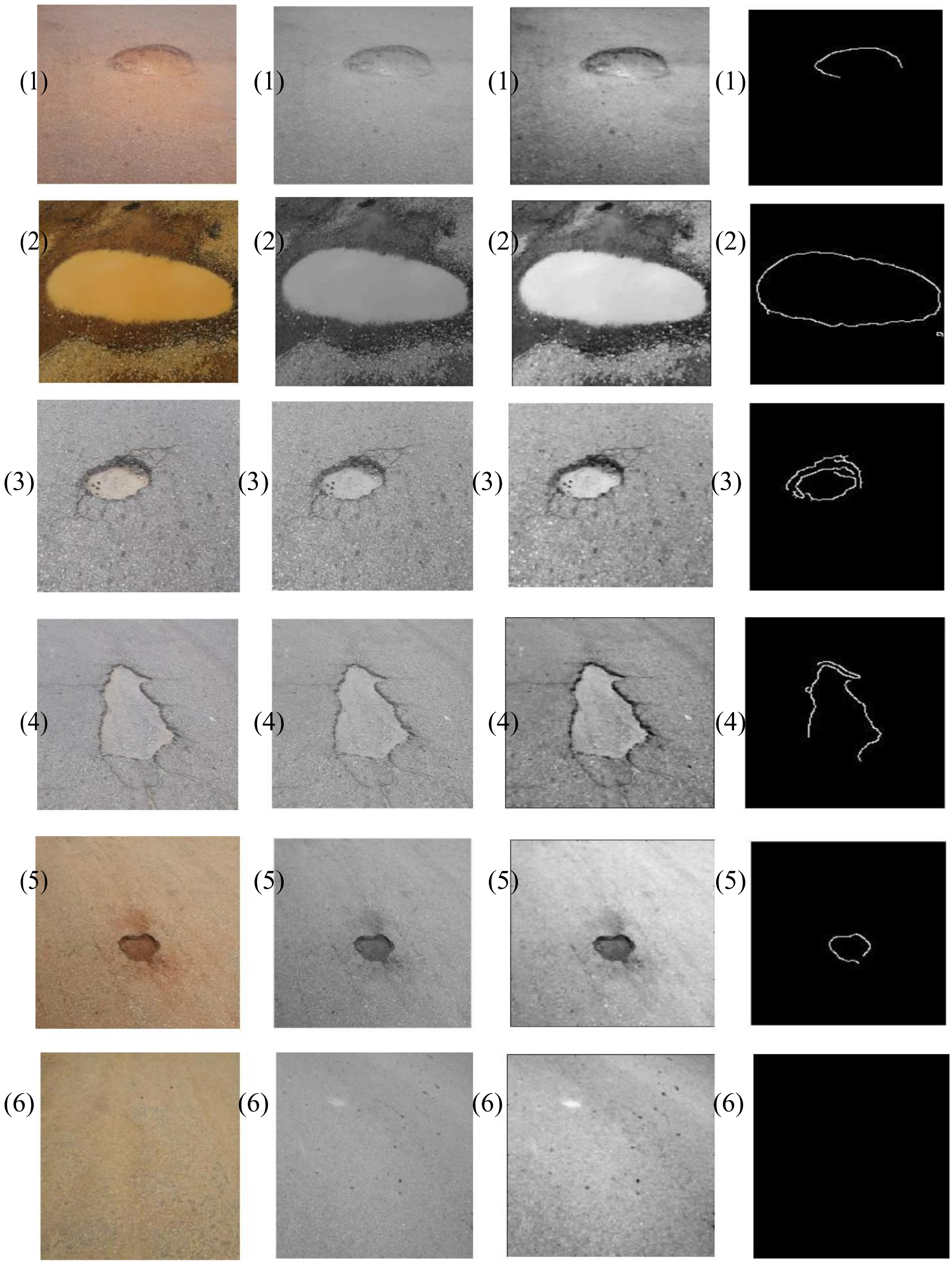
$$\text{FPR} = 1 - \frac{TN}{FP+TN} \qquad (3.7)$$

CHAPTER FOUR

4.0 RESULTS AND DISCUSSION

4.1 Results of Image Pre-Processing and Segmentation

Figure 4.1, shows some of the road surface images captured at different light illuminations, alongside the individual result obtained after image pre-processing and segmentation of the images. Figure 4.1(a), presents typical scenario of input raw data, which were denoised using the median filter and the corresponding outputs (see Figure 4.1(b)) was fed into the discrete wavelet transform. It can be observed that the blurriness in the median filtered images (see Figure 4.1(b)) was improved upon and the edges of the pothole anomaly become sharper as seen in the output of the DWT shown in Figure 4c. However, the output of the segmented images using the canny edge extractor was able to extract edges of portion of the images with the road anomaly from the background as shown in Figure 4.1(d). This was made possible by the utilization of discrete wavelet transform that helps preserve and enhance the edges of the processed image towards improving the performance of the CNN for accurate detection and classification.



(a) Original images (b) Median Images (c) DWT images (d) Segmentation Figure

4.1: Image pre-processing and segmentation result of captured road images.

4.2 The Detection and Classification Result

The binary classification result obtained from our algorithm is shown in the confusion matrix in Figure 4.2. The confusion matrix shows the number of True Positives (TP), True Negative (TN), False Positives (FP), and False Negatives (FN) result from the model.

57 5

3 55

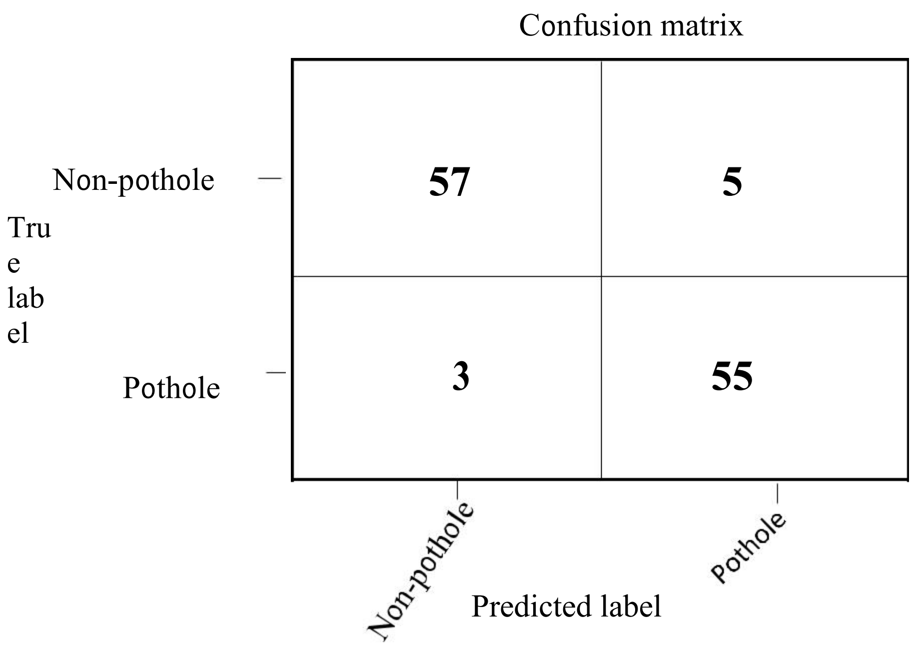


Figure 4.2: Confusion Matrix Result

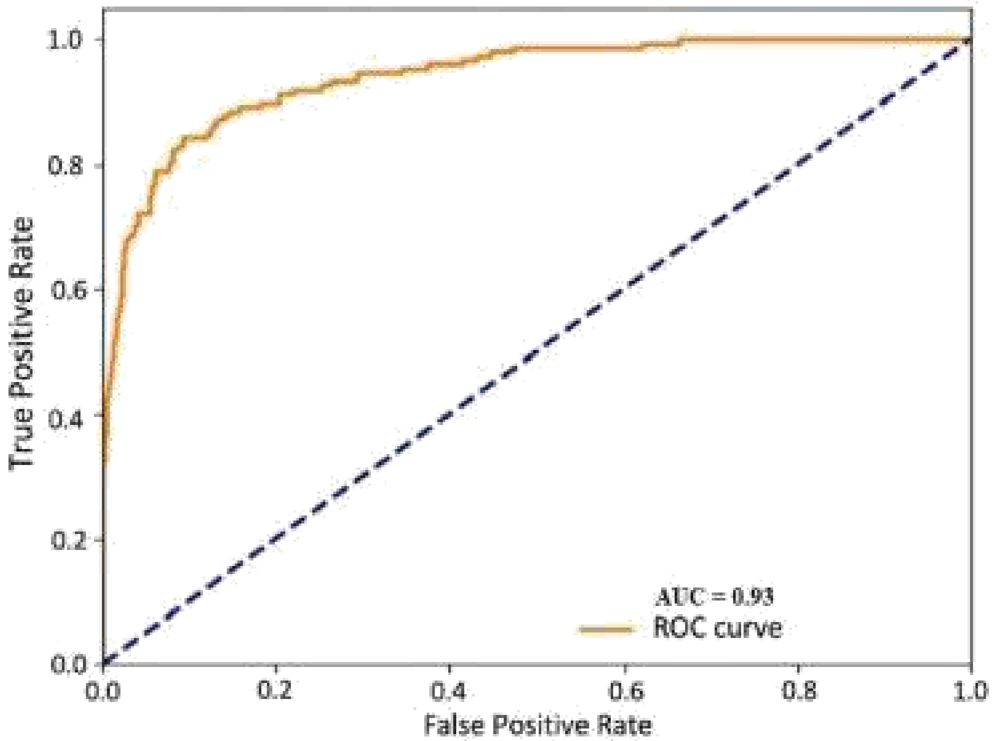


Figure 4.3: Receiver Operating Characteristic (ROC) Result

The ROC curve in Figure 4.3 shows the trade-off between the sensitivity (TPR) and specificity (1-FPR) of our classifier model (CNN). The ROC curve was plotted by the MATLAB prototype that simulated the classifier model, by making use of the outcomes of True Positive Rate and False Positive Rate outcomes from the confusion matrix, and result of the Area under ROC curve (AUC) obtained was 0.93, which was shown on the graph. This result means that the probability of the classifier, classifying either a pothole or smooth road images accurately, is higher than the probability of false classification. AUC tells how much a model is capable of distinguishing between classes. The higher the AUC, the better the model is at predicting pothole's images as pothole and smooth road images as smooth road.

4.3 Results of Performance Evaluation

Result obtained from the confusion matrix is as follows:

True positive (TP): 55

True Negative (TN): 57

False Positive (FP): 5

False Negative (FN): 3

In order to evaluate the performance of this research work, makes use of equation (3.4), (3.5) and (3.6).

Therefore,

$$\text{Accuracy} = \frac{55+57}{55+57+5+3}$$

$$= \frac{112}{120} = 0.9333$$

$$= 0.9333 = 93.33\%$$

$$\text{Precision} = \frac{55}{55+5}$$

$$= \frac{55}{60} = 0.9167$$

$$= 0.9167 = 91.67\%$$

$$\text{Recall} = \frac{55}{55+3}$$

$$= \frac{55}{58} = 0.9483$$

$$= 0.9483 = 94.83\%$$

This research work has an accuracy of 93.33%, precision of 91.67%, and recall of 94.83%. Figure 4.4, 4.5 and 4.6, shows chart comparisons of this research work to other 2D vision-based techniques in terms of their accuracy, precision and recall.

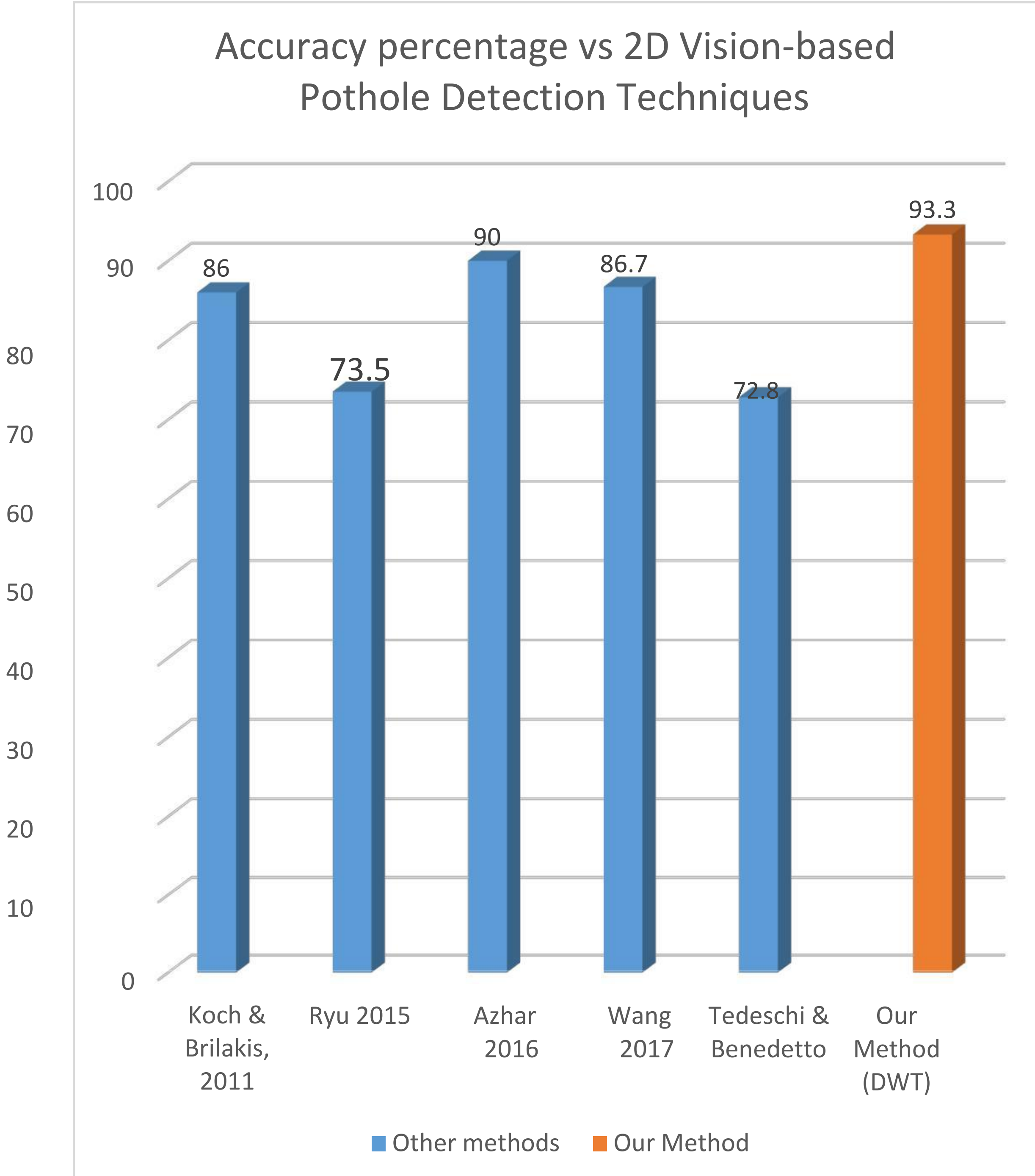


Figure 4.4: Accuracy Comparison with other 2D Vision-Based Techniques.

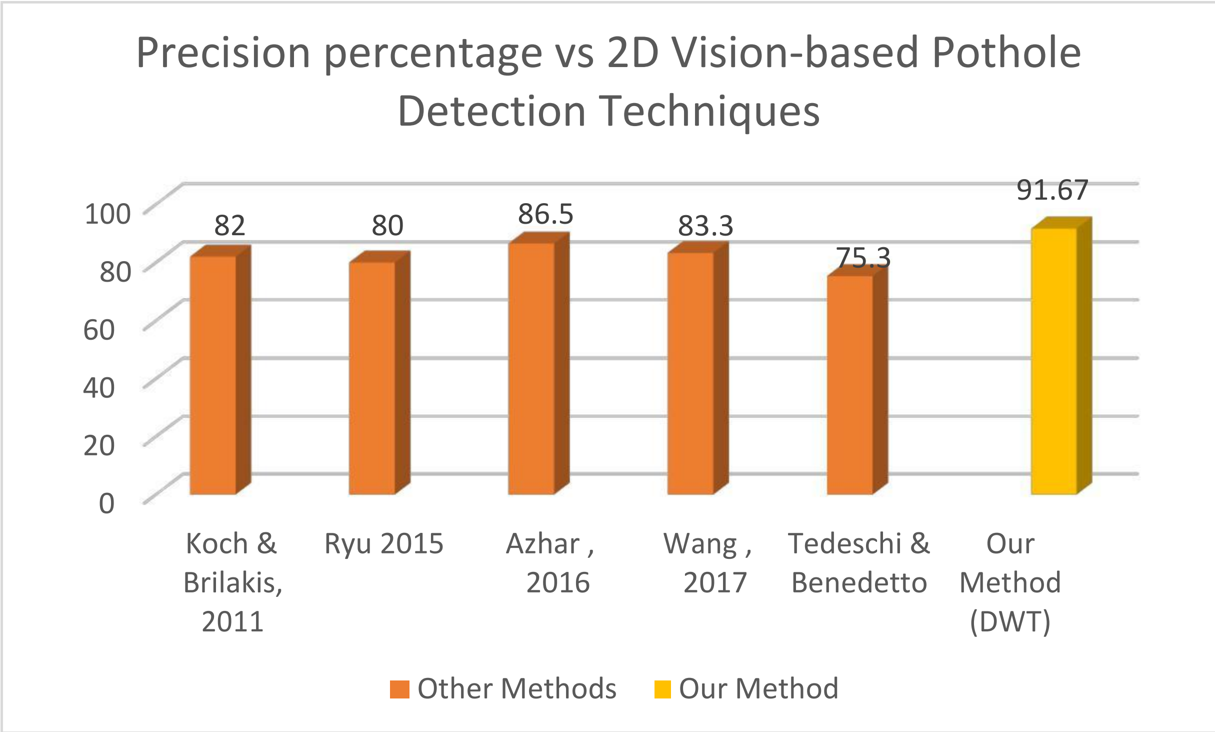


Figure 4.5: Precision Comparison with other 2D Vision-Based Techniques.

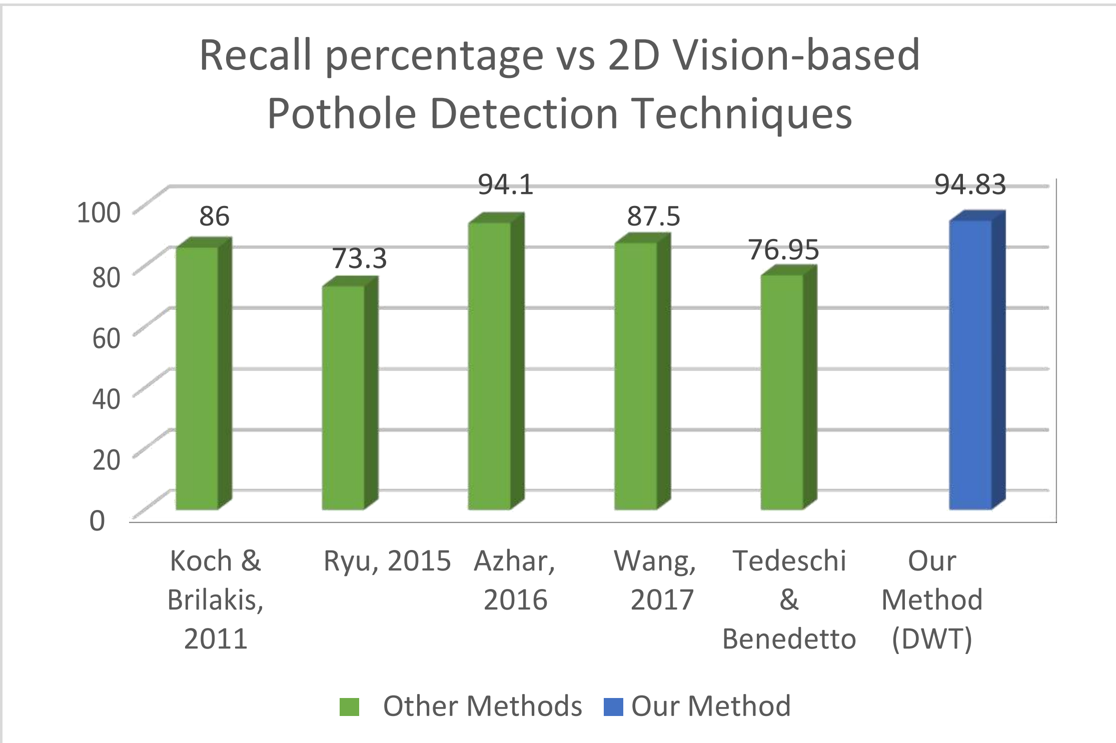


Figure 4.6: Recall Comparison with other 2D Vision-Based Techniques.

4.4 Discussion of Result

The accuracy of this research work is 93.33%, with precision of 91.67%, and recall of 94.83%. In other works reported in literature, different features are adopted for use in the detection of road defects in road images. Some of the features include, the geometric properties in the image, size, ellipticity, linearity, compactness, gradients, edge extraction/detection, among others. All these features are been used independently, and taken into consideration to actualize the detection of road defect by different researchers in the vision-based detection methods. This research made use of edge detection in actualizing pothole detection.

The method proposed by Ryu *et al*. (2015) and Koch & Brilakis (2011) made use of geometric criterion calculation and also the extraction of texture information to identify potholes in the images. However, the detection accuracy in these methods were low (73.5% and 86% respectively), which made Koch & Brilakis (2011) to recommend the use of machine learning for detection for future works, which was adopted in the research work.

Wang *et al*. (2017) also made use of geometric criterion to judge if a road images has a pothole or not, by using Wavelet field and Markov random field for segmentation. However, their proposed method also had a low detection accuracy (86.7%) compared to this research, and the use of Discrete Wavelet Transform along Markov random field was proposed for future works.

Also, Histograms of oriented gradients feature was adopted by Azhar *et al*. (2016) .They also made use of a machine learning model for detection (Naïve Bayes classifier) and the detection accuracy of the algorithm was 90%. However, few dataset were used for training and testing of the algorithm (120 road images), compared to this research (400 road images).

This research is able to detect different shapes and sizes of potholes, it is also able to segment and detect potholes filled with water during foggy weather conditions as seen in Figure 4.0(2). Therefore, it can be said that, this research developed a model for pothole detection with an improved detection accuracy, precision and recall than Koch & Brilakis (2011) and Wang *et al*. (2017) and other 2D-vision based pothole detection techniques presented in literature. However, during simulation, it took approximately two (2) seconds to pre-process and segment an input road image, while it took about fifteen (15) seconds to detect and classify the road image, which is not achievable for real-time detection.

CHAPTER FIVE

5.0 CONCLUSION AND RECOMMENDATION

5.1 Conclusion

A 2D vison-based detection approach aim towards improving road anomaly (pothole) detection on asphalt roads is presented in this research work. This entailed capturing different road surface images to form a dataset for use, pre-processing these images using median filter and discrete wavelet transform (DWT) for denoising, preservation and enhancement of edges of the acquired images. A canny edge extractor was used for segmentation of the pre-processed images, which was subsequently fed into a ResNet 50 pre-trained CNN Model for detection and classification of the road images. This research was simulated on MATLAB prototype, and results obtained from the simulation shows that, the developed model is an improvement to the existing 2D vision-based road anomaly (pothole) detection techniques proposed in literatures, having a detection accuracy of 93.3%, precision of 91.67%, and recall of 94.83%.

5.2 Recommendations

Future works should explore:

I. Improving the detection / classification processing time of this model, towards real-time potholes anomaly detection.

AI. A dynamic segmentation threshold should be explored in order to get better segmentation results for detection, thus improving the detection accuracy of this research.

5.3 Contribution to Knowledge

From literatures, a lot of vision-based pothole detection algorithms are been proposed each having their strengths and limitations. However, the most prominent / common limitation faced in most of the literatures is the inability for the algorithms to detect potholes accurately in foggy and light illumination weather conditions. Therefore, this research was focused on the implementation of a new vision-based pothole detection algorithm with better detection accuracy even during foggy and light illumination conditions.

REFERENCES

Akagic, A., Buza, E., & Omanovic, S. (2017). Pothole detection: An efficient vision-based method using rgb color space image segmentation. In *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 1104-1109). IEEE, https://doi.org/10.23919/MIPRO.2017.7973589

Akagic, A., Buza, E., Omanovic, S., & Karabegovic, A. (2018). Pavement crack detection using Otsu thresholding for image segmentation. In *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 1092-1097). IEEE, https://doi.org/10.23919/MIPRO.2018.8400199

Akarsu, B., Karaköse, M., Parlak, K.., Akin, E., & Sarimaden, A. (2016). A fast and adaptive road defect detection approach using computer vision with real time implementation. *International Journal of Applied Mathematics Electronics and Computers*, (Special Issue-1), 290-295, https://doi.org/10.18100/ijamec.270546

Aldagheiri, M. (2009). The role of the transport road network in the economic development of Saudi Arabia. *WIT Transactions on the Built Environment*, *107*, 275-286, https://doi.org/10.2495/UT090251

Aparna, Bhatia, Y., Rai, R., Gupta, V., Aggarwal, N., & Akula, A. (2019). Convolutional neural networks-based potholes detection using thermal imaging. *Journal of King Saud University-Computer and Information Sciences*, https://doi.org/10.1016/j.jksuci.2019.02.004

Azhar, K., Murtaza, F., Yousaf, M. H., & Habib, H. A. (2016). Computer vision-based detection and localization of potholes in asphalt pavement images. In *2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)* (pp. 1-5). IEEE, https://doi.org/10.1109/CCECE.2016.7726722

Bello-Salau, H., Aibinu, A. M., Onwuka, E. N., Dukiya, J. J., & Onumanyi, A. J. (2014). Image processing techniques for automated road defect detection: A survey. In *2014 11th International Conference on Electronics, Computer and Computation (ICECCO)* (pp. 1-4). IEEE, https://doi.org/10.1109/ICECCO.2014.6997556

Bello-Salau, I. H., Onumanyi, A. J., Aibinu, A. M., Onwuka, E. N., Dukiya, J. J., & Ohize, H. (2019). A survey of accelerometer-based techniques for road anomalies detection and characterization. *International Journal of Engineering Science and Application*, *3*(1), 8-20.

Bhatnagar, A. (2017). Image Segmentation Using Canny Edge Detection Technique Kirti. vol, 4, 8-14

Bubenikova, E., & Halgas, J. (2012). Application of image processing in intelligent transport systems. IFAC Proceedings Volumes, 45(7), 53-56, https://doi.org/10.3182/20120523-3-CZ-3015.00012

Buchinger, D., & Silva, A. G. (2014). Anomalies detection in asphalt pavements: a morphological image processing approach. Revista Brasileira de Computação Aplicada, 6(1), 121-129, https://doi.org/10.5335/rbca.2014.3661

Buza, E., Akagic, A., Omanovic, S., & Hasic, H. (2017). Unsupervised method for detection of high severity distresses on asphalt pavements. In 2017 IEEE 14th International Scientific Conference on Informatics (pp. 45-50). IEEE. https://doi.org/10.1109/INFORMATICS.2017.8327220

Buza, E., Omanovic, S., & Huseinovic, A. (2013). Pothole detection with image processing and spectral clustering. In Proceedings of the 2nd International Conference on Information Technology and Computer Networks (Vol. 810, p. 4853).

Byun, Y., Song, H., & Baek, K. (2018). Pothole Detection using Machine Learning. Proceedings of the Advance Science and Technology (pp. 151-155), https://doi.org/10.14257/astl.2018.150.35

Chambon, S., & Moliard, J. M. (2011). Automatic road pavement assessment with image processing: review and comparison. International Journal of Geophysics (Vol. 2011). https://doi.org/10.1155/2011/989354

Deng, L., & Yu, D. (2014). Deep learning: methods and applications. Foundations and trends in signal processing, 7(3–4), 197-387. https://doi.org/10.1561/2000000039

Erkan, U., Gokrem, L., & Enginoglu, S. (2018). Different applied median filter in salt and pepper noise. *Computers & Electrical Engineering*, *70*, 789-798.

Fan, S., & Chan-Kang, C. (2005). *Road development, economic growth, and poverty reduction in China*. Retrieved from https://books.google.com/books?hl=en&lr=&id=f3c-l4qIYtoC&oi=fnd&pg=PR5&dq=road+reducing+poverty&ots=nNX-QyNmAO&sig=PnX8Gkv_eqMj9e5LRzXmWbaSMrQ

George, G., Oommen, R. M., Shelly, S., Philipose, S. S., & Varghese, A. M. (2018). A survey on various median filtering techniques for removal of impulse noise from digital image. In 2018 conference on emerging devices and smart systems (ICEDSS) (pp. 235-238). IEEE. https://doi.org/10.1109/ICEDSS.2018.8544273

Google. (2020). Google Maps directions of Eastern by-pass Tundu wada road, Minna, Niger state, Nigeria. 9°36'40.4"N6°34'38.6"E. Retrived Febuary 16th, 2020, from https://goo.gl/maps/RP1eRY4app4DuEacd8

Gulcu, A., & Kus, Z. (2019). A Survey of Hyper-parameter Optimization Methods in Convolutional Neural Networks. *7*(2), 503–522.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016–Decem*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Hegde, S., Mekali, H., & Varaprasad, G. (2014). Pothole detection and inter vehicular communication. *2014 IEEE International Conference on Vehicular Electronics and Safety, ICVES 2014*, (March), 84–87. https://doi.org/10.1109/ICVES.2014.7063729

Hoang, N. D., Nguyen, Q. L., & Tien Bui, D. (2018). Image processing–based classification of asphalt pavement cracks using support vector machine optimized by artificial bee colony. *Journal of Computing in Civil Engineering*, *32*(5), 04018037. https://doi.org/10.1061/(asce)cp.1943-5487.0000781

Hoang, N. D. (2018). An Artificial Intelligence Method for Asphalt Pavement Pothole Detection Using Least Squares Support Vector Machine and Neural Network with Steerable Filter-Based Feature Extraction. *Advances in Civil Engineering*, *2018*. https://doi.org/10.1155/2018/7419058

Huidrom, L., Das, L. K., & Sud, S. K. (2013). Method for Automated Assessment of Potholes, Cracks and Patches from Road Surface Video Clips. *Procedia - Social and Behavioral Sciences*, *104*, 312–321. https://doi.org/10.1016/j.sbspro.2013.11.124

Jahanshahi, M. R., & Masri, S. F. (2012). Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures. *Automation in Construction*, *22*, 567–576. https://doi.org/10.1016/j.autcon.2011.11.018

Jensen, J. R., & Lulla, K. (1987). Introductory digital image processing: A remote sensing perspective. *Geocarto International*, *2*(1), 65. https://doi.org/10.1080/10106048709354084

Kabade, A. L., & Sangam, V. G. (2016). Canny edge detection algorithm. *2nd International Journal of Advance Research in Electronics and Communication Engineering (IJARECE)*, *5*(5), 1292–1295. Retrieved from http://dasl.mem.drexel.edu/alumni/bGreen/www.pages.drexel.edu/_weg22/can_tut.html

Khan, A., Sohail, A., Zahoora, U., & Qureshi, A. S. (2019). A survey of the recent architectures of deep convolutional neural networks. Artificial Intelligence Review, 53(8), 5455-5516.

Khan, S., Rahmani, H., Shah, S. A. A., & Bennamoun, M. (2018). A guide to convolutional neural networks for computer vision. Synthesis Lectures on Computer Vision, 8(1), 1-207. https://doi.org/10.2200/s00822ed1v01y201712cov015

Kim, T., & Ryu, S. K. (2014). Review and analysis of pothole detection methods. Journal of Emerging Trends in Computing and Information Sciences, 5(8), 603-608. Retrieved from http://www.cisjournal.org

Koch, C., & Brilakis, I. (2011). Pothole detection in asphalt pavement images. Advanced Engineering Informatics, 25(3), 507-515. https://doi.org/10.1016/j.aei.2011.01.002

Lee, S., Kim, S., An, K. E., Ryu, S. K., & Seo, D. (2018). Image processing-based pothole detecting system for driving environment. *2018 IEEE International Conference on Consumer Electronics, ICCE 2018*, *2018–Janua*, 1–2.

https://doi.org/10.1109/ICCE.2018.8326141

Lemer, A. C. (1999). Building public works infrastructure management systems for achieving high return on public assets. Public Works Management & Policy, 3(3), 255-272. https://doi.org/10.1177/1087724X9933007

Liu, X., Ai, Y., & Scherer, S. (2017). Robust image-based crack detection in concrete structure using multi-scale enhancement and visual features. In 2017 IEEE International Conference on Image Processing (ICIP) (pp. 2304-2308). IEEE. https://doi.org/10.1109/ICIP.2017.8296693

Medina, R., Llamas, J., Zalama, E., & Gomez-Garcia-Bermejo, J. (2014). Enhanced automatic detection of road surface cracks by combining 2D/3D image processing techniques. *2014 IEEE International Conference on Image Processing, ICIP 2014*, 778–782. https://doi.org/10.1109/ICIP.2014.7025156

Nienaber, S., Booysen, M., & Kroon, R. (2015). Detecting Potholes Using Simple Image Processing Techniques and Real-World Footage. *Proceedings of the 34th Southern African Transport Conference (SATC*, (Satc), 153–164. Retrieved from https://www.semanticscholar.org/paper/Detecting-potholes-using-simple-image-processingNienaberBooysen/9375f4d7fb035924c709c58841b3f5acb7b989aa%0A

Onoyan-Usina, A., Lazhi, Y. G., & Itomi-Ushi, D. P. (2013). Bad drainage and its effects on road pavement conditions in Nigeria. Civil and Environmental Research, 3(10), 7-15. Retrieved from https://www.cabdirect.org/cabdirect/abstract/20133422133

Ouma, Y. O., & Hahn, M. (2016). Wavelet-morphology based detection of incipient linear cracks in asphalt pavements from RGB camera imagery and classification using circular Radon transform. *Advanced Engineering Informatics*, *30*(3), 481-499. https://doi.org/10.1016/j.aei.2016.06.003

Ouma, Y. O., & Hahn, M. (2017). Pothole detection on asphalt pavements from 2D-colour pothole images using fuzzy c-means clustering and morphological reconstruction. Automation in Construction, 83, 196-211. https://doi.org/10.1016/j.autcon.2017.08.017

Radopoulou, S. C., & Brilakis, I. (2015). Patch detection for pavement assessment. Automation in Construction, 53, 95-104. https://doi.org/10.1016/j.autcon.2015.03.010

Ryu, S. K., Kim, T., & Kim, Y. R. (2015). Image-based pothole detection system for ITS service and road management system. Mathematical Problems in Engineering, vol. 2015, (pp. 1-10). https://doi.org/10.1155/2015/968361

Sharma, M., Sharma, M. K., Leeprechanon, N., Anotaipaiboon, W., & Chaiyasarn, K. (2017). Digital image patch based randomized crack detection in concrete structures. In 2017 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-7). IEEE. https://doi.org/10.1109/ICCCI.2017.8117742

Siriborvornratanakul, T. (2018). An automatic road distress visual inspection system using an onboard in-car camera. Advances in Multimedia, vol. 2018, (pp. 1-10). https://doi.org/10.1155/2018/2561953

Smith, A. (2010). Image segmentation scale parameter optimization and land cover classification using the Random Forest algorithm. Journal of Spatial Science, 55(1), 69-79. https://doi.org/10.1080/14498596.2010.487851

Smith, D. S., & Basson, I. J. (2006). Shape and distribution analysis of Merensky Reef potholing, Northam Platinum Mine, Western Bushveld Complex: implications for pothole formation and growth. Mineralium Deposita, 41(3), 281-295. https://doi.org/10.1007/s00126-006-0059-5

Stork, M. (2013). Median filters theory and applications. In Third International Conference on Electrical and Electronics Engineering Papers-Chamber of Electrical Engineering, Bursa, Turkey (No. 3, pp. 1-5).

Sultani, W., Mokhtari, S., & Yun, H. B. (2017). Automatic pavement object detection using superpixel segmentation combined with conditional random field. IEEE Transactions on Intelligent Transportation Systems, 19(7), 2076-2085.

Tedeschi, A., & Benedetto, F. (2017). A real-time automatic pavement crack and pothole recognition system for mobile Android-based devices. Advanced Engineering Informatics, 32, 11-25. https://doi.org/10.1016/j.aei.2016.12.004

Tsai, Y. C., & Chatterjee, A. (2018). Pothole detection and classification using 3D technology and watershed method. Journal of Computing in Civil Engineering, 32(2), 04017078. https://doi.org/10.1061/(asce)cp.1943-5487.0000726

Tunde, A. M., & Adeniyi, E. E. (2012). Impact of road transport on agricultural development: a Nigerian example. *Ethiopian Journal of Environmental Studies and Management*, *5*(3), 232-238.

Vosoughi, A., Shamsollahi, M. B., & Vosoughi, A. (2009). Nonsubsampled higher-density discrete wavelet transform: Filter design and application in image contrast enhancement. In 2009 16th IEEE International Conference on Image Processing (ICIP) (pp. 3165-3168). IEEE.

Van de Walle, D. (2002). Choosing rural road investments to help reduce poverty. World development, 30(4), 575-589. https://doi.org/10.1596/1813-9450-2458

Wang, L., Chen, Z., & Wu, J. (2017). An opportunistic routing for data forwarding based on vehicle mobility association in vehicular ad hoc networks. Information, 8(4), 140. https://doi.org/10.3390/info8040140

Wang, P., Hu, Y., Dai, Y., & Tian, M. (2017). Asphalt pavement pothole detection and segmentation based on wavelet energy field. Mathematical Problems in Engineering, 2017. https://doi.org/10.1155/2017/1604130

Williams, J. R., & Amaratunga, K. (1994). Introduction to wavelets in engineering. International journal for numerical methods in engineering, 37(14), 2365-2388. https://doi.org/10.1002/nme.1620371403

Xia, W. (2018). An approach for extracting road pavement disease from HD camera videos by deep convolutional networks. In 2018 International Conference on Audio, Language and Image Processing (ICALIP) (pp. 418-422). IEEE. https://doi.org/10.1109/ICALIP.2018.8455601

Yadav, V., & Bethard, S. (2019). A survey on recent advances in named entity recognition from deep learning models. Proceedings of the 27th International Conference on Computational Linguistics, Santa Fe, 2018, 2145–2158. Retrieved from http://arxiv.org/abs/1910.11470

Zhang, D., Li, Q., Chen, Y., Cao, M., He, L., & Zhang, B. (2017). An efficient and reliable coarse-to-fine approach for asphalt pavement crack detection. Image and Vision Computing, 57, 130-146. https://doi.org/10.1016/j.imavis.2016.11.018

Zhang, Z., Ai, X., Chan, C. K., & Dahnoun, N. (2014). An efficient algorithm for pothole detection using stereo vision. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 564-568). IEEE. https://doi.org/10.1109/ICASSP.2014.6853659

Mednis, A., Strazdins, G., Zviedris, R., Kanonirs, G., & Selavo, L. (2011). Real time pothole detection using android smartphones with accelerometers. In 2011 International conference on distributed computing in sensor systems and workshops (DCOSS) (pp. 1-6). IEEE. https://doi.org/10.1109/DCOSS.2011.5982206

540 killed in road crashes in January - FRSC - Premium Times Nigeria. (2019). Retrieved January 10, 2020, from https://www.premiumtimesng.com/news/headlines/328229-540-killed-in-road-crashes in-january-frsc.html

APPENDIX A

Matlab Code for the entire Research Work

The entire code was run on MATLAB prototype.

```
%------------------ pre-processing stage-----------------------
img = imread('IMG_20190621_143211.jpg'); %input
  image B = imresize(img, [400 300]); %resizing the image
  G = rgb2gray(B); %convertion of image into grayscale
  image Medianfilter = medfilt2(G); %applying Median filter
%imshow(Medianfilter);

  colormap gray
  [LoD, HiD] = wfilters('haar','d'); %applying Discrete wavelet transform
  [cA,cH,cV,cD] = dwt2(Medianfilter,LoD,HiD, 'mode', 'symh');
  imagesc(cA)
%---------------------------------------------------------------

%--------------- Segmentation stage-----------------------------
  colormap gray
  title('vertical')
   h = edge(cA, 'canny', 0.75); %applying canny edge detector
 imshow (h);
 %se = ones(30,30);
 %IM2 = imdilate(h,se);
```

```
%Figure
%imshow (IM2);
imwrite (h, 'output.jpg', 'jpg');
%-----------------------------------------------------------

%-------------------------CNN------------------------------
outputFolder = fullfile('matlab');
rootFolder = fullfile(outputFolder, 'imageP');
%------------------classification of Road images----------
categories = {'pothole', 'road'};

imds = imageDatastore(fullfile(rootFolder,categories),'LabelSource','foldernames');

tbl = countEachLabel(imds);
minSetCount = min(tbl{:,2});

imds = splitEachLabel (imds, minSetCount, 'randomize' );
countEachLabel(imds);

pothole = find(imds.Labels == 'pothole', 1);
road = find(imds.Labels == 'road', 1);

net = resnet50(); %Loading Resnet50 model
```

```
%---------------- Training of the Model----------------------

[trainingSet, testSet] = splitEachLabel(imds, 0.3, 'randomize');


imageSize = net.Layers(1).InputSize;


augmentedTrainingSet = augmentedImageDatastore(imageSize, ...

    trainingSet, 'colorPreprocessing', 'gray2rgb');


augmentedTestSet = augmentedImageDatastore(imageSize,...

    testSet, 'colorPreprocessing', 'gray2rgb');


%%%%  Getting the weight of second Convolation

Layers w1 = net.Layers(2).Weights;

%%converting the matric to

image w1 = mat2gray(w1);


%%       Figure

%%       montage(w1)

%%       tile ('First Convolutional Layer Weight')


featureLayer = 'fc1000';

trainingFeatures = activations(net, ...

    augmentedTrainingSet, featureLayer, 'MiniBatchsize', 32, 'OutputAS', 'columns');
```

```
trainingLables = trainingSet.Labels;
classifier = fitcecoc(trainingFeatures, trainingLables, ...
    'Learner', 'Linear', 'Coding', 'onevsall', 'ObservationsIn', 'columns');

testFeatures = activations(net, ...
    augmentedTestSet, featureLayer, 'MiniBatchsize', 32, 'OutputAS', 'columns');

predictLabels = predict(classifier, testFeatures, 'ObservationsIn', 'columns');

testLabels = testSet.Labels;
confMat = confusionmat(testLabels,predictLabels );
confMat = bsxfun(@rdivide, confMat, sum(confMat,2));

mean(diag(confMat));
%---------------------------------------------------------
%-----Detection and Classification of the processed road image--
newImage = imread(fullfile ('output.jpg'));

ds = augmentedImageDatastore(imageSize,...
    newImage, 'colorPreprocessing', 'gray2rgb');

imageFeatures = activations(net, ...
    ds, featureLayer, 'MiniBatchsize', 32, 'OutputAS', 'columns');

label = predict(classifier, imageFeatures, 'ObservationsIn', 'columns');
```

```
sprintf('The loaded image belongs to %s class', label) %display output

%--------------------------------------------------------------------------
```

APPENDIX B

Summary of Reviewed Papers in Literature

Table 2.1: Summary of related works.

AUTHOR	TECHNIQUES	STRENGTH	LIMITATION	ACCURACY	PRECISION	RECALL
Akarsu *et al*. (2016)	Using Binary transform and Morphological operation	High accuracy rates in different types of roads Ability to classify errors on different road surface. Real time detection system which classify errors on different road surface. Cracks/ defect thinner than 1mm can be detected.	Road color can affect the accuracy rates of the operation.	_	_	_
Wang *et al*. (2017)	Wavelet Energy field model and Markor random field model.	Cost effective in implementation. Good pothole segmentation results for different kinds of potholes.	Complexity. High Processing time Detection accuracy is by light illumination conditions.	86.7%	83.3%	87.5%

Akagic *et al*. (2017)	RGB color space Image.	Less complex in implementation. Low computational cost. Real time implementation.	Detection accuracy is affected by light illumination conditions. The effectiveness of the method depends on the extraction of ROI accuracy.	82%	_	_
Koch & Brilakis (2011)	Morphological thinning and Elliptic regression.	It has good detection accuracy. Simplicity.	Detection accuracy is affected by light illumination conditions.	86%	82%	86%
Buchinger & Silva (2014)	Morphological image processing	Simplicity. Able to detect cracks, but detect potholes more accurately.	Detection accuracy is affected by light illumination conditions.	_	_	_
Liu *et al*. (2017)	Multi-scale enhancement and visual features	Reduces the rate of false detection of road defect.	Complexity.	_	_	_
Hoang *et al*. (2018)	Artificial Neural Network (ANN), and Least squares support vector machine (LS-SVM)	The two AI combined as higher detection accuracy.	Complexity. High Processing time. Cannot estimate the size of potholes.	LS-SVM= 88.75% ANN= 85.25%	-	-
Huidrom *et al*. (2013)	DFS algorithm for segmentation,	The use of CDDMA algorithm can detect and measure pothole, cracks and patches distresses	Complexity.	-	-	-

	and CDDMC algorithm for detection.	efficiently and accurately in one pass.				
Ryu *et al*. (2015)	Morphological operations, Histogram Shaped-Based Thresholding, and Ordered Histogram Intersection (OHI)	Average processing time	Complexity. Not applicable for real time implementation Varying accuracy under different weather conditions.	73.5%	80%	73.3%
Azhar *et al*. (2016)	Histograms of oriented gradients (HOG) features.	Low-cost and computationally efficient. Average processing time	Not applicable for real time implementation	90%	86.5%	94.1%
Radopoulou & Brilakis (2015)	Histogram equalization for image enhancement, and Standard deviation of gray-level intensity.	Fast processing time. Low implementation cost.	Detection accuracy rate is affected by light and weather conditions. Cannot detect other type of road defects	75%	82%	86%
Z. Zhang *et al*. (2014)	Stereo vision	More accurate result than 2D image techniques. The ROI is easily determined.	Complexity. High computation cost. High processing time.	-	-	-

			False detection also occurs, due to error in the disparity calculations.			
Sultani *et al*. (2017)	Superpixel Segmentation and Conditional Random Field for detection.	High efficiency. Cost effective. Can be used to detect any object on roads.	Complexity. Higher processing time	-	-	-
Buza *et al*. (2018)	Binary image and Simple Linear Iterative Clustering (SLIC) superpixels method for image segmentation	Low computational pre-processing steps. Good processing time.	Detection accuracy is affected by light illumination and weather conditions. Does not detect smaller potholes with diameter less than 100 mm accurately	-	-	-
Sharma *et al*. (2017)	Random forests classifier for Crack classification.	Detect crack efficiently.	High processing time High computational cost	-	-	-
Tedeschi & Benedetto (2017)	Local binary pattern (LBP) cascade classifier for training and classifying road distress	Low computational cost. Real time implementation	Low detection accuracy compared to other methods. Cannot extract crack curves or pothole shapes.	72.8%	75.35%	76.95%
Medina *et al*. (2014)	Combination of 2D/3D image processing	Detect cracks in the pavement of the road combining visual appearance and geometrical	Complexity. High implementation cost.			

	techniques.	information.	High processing time. Cannot be implemented on real time	95.25%	-	-
Lee *et al*. (2018)	Wavelet energy field and Superpixel.	Simplicity.	False detection of 35%.	-	-	-
Ouma & Hahn (2017)	Fuzzy c-means clustering and morphological reconstruction.	Cost effective.	Complexity. Detection accuracy is affected by light illumination and weather conditions	87.5%	-	-
Buza *et al*. (2013)	Spectral clustering	Cost effective. It proposed method can identify potholes and estimate their surface	High processing time.	81%	-	-
Tsai & Chatterjee (2017)	Detection and classification potholes in roads using 3D technology and watershed method.	Severity levels of each potholes can be assessed. High detection accuracy.	Complexity. High implementation cost. Not applicable for real time implementation.	94.97%	90.80%	98.75%

APPENDIX C

Publications

Oyinbo, A. M., Mohammed, A. S., Zubair, S., & Michael, E. (2021). A Review of Different Proposed Image Detection Techniques for Road Anomalies Detection. *Proceeding of the 1st NSE Minna Branch Engineering Conference,* 113–119.

Oyinbo, A. M., Salau, H. B., Mohammed, A. S., Zubair, S., Adejo, A., & Abdulkarim, H. T. (2020). *Towards an Improved Potholes Anomaly Detection Based on Discrete Wavelet Transform and Convolution Neural Network : A Proposal*. Nigerian Journal of Engineering, *27*(2), 86–91.

Bello-Salau, H., Onumanyi, A. J., Salawudeen, A. T., Mu'azu, M. B., & Oyinbo, A. M. (2019). An Examination of Different Vision based Approaches for Road Anomaly Detection. In *2019 2nd International Conference of the IEEE Nigeria Computer Chapter (NigeriaComputConf)* (pp. 1-6). IEEE.